TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

JINDOU CHEN

INTERPOLATION AND EXTRAPOLATION METHODS FOR WLAN-BASED POSITIONING

Master of Science thesis

Examiners: Assoc.Prof.
Elena-Simona Lohan
Prof. Pertti Järventausta

Examiner and topic approved by the Faculty Council of the Faculty of Computer and Electrical Engineering on
November 5th, 2014

ABSTRACT

JINDOU CHEN: INTERPOLATION AND EXTRAPOLATION METHODS FOR WLAN-BASED POSITIONING

Tampere University of technology
Master of Science Thesis, 57 pages
February 2015
Master's Degree Programme in Electrical Engineering
Major: Smart Grids
Examiners: Assoc.Prof Elena-Simona Lohan
Prof. Pertti Järventausta

Keywords: Delaunay Triangulation, Interpolation, Extrapolation, Wireless local area networks (WLANs), Mean absolute error.

WLAN-based positioning is obtaining more and more attention in the research field nowadays. In order to create better Location Based Services (LBSs), the demand to achieve higher user location accuracy is increasing. This thesis aims at studying the effect of different interpolation and extrapolation methods in the WLAN-based indoor positioning, based on collected WLAN data.

Depending on the embraced positioning method, there are various errors in WLAN-based positioning, such as calibration error, measurement errors, shadowing, etc. The motivation of this work came from trying to decrease the positioning error in the absence of complete information about the indoor environment. This can be done by using interpolation and extrapolation methods, which are widely used in image processing nowadays. However, they are also an available and efficient way to deal with WLAN-based positioning studies.

Among interpolation methods, Delaunay triangulation can partly avoid introducing distortions in the measurement databases. Therefore, it makes sense to investigate triangulation based methods and to study their usefulness in the WLAN context. Practically, it is very hard to extrapolate appropriately and the implementation of the extrapolation is much more complex than the one of the interpolation. Thus in this thesis, simple extrapolation methods have been performed.

The results here are based on measurement data. The performance of each method is analyzed in terms of the error between the received signal strengths (RSS) coming from the measurements and the RSS obtained through interpolation and extrapolation. WLAN data was collected from several buildings of Tampere University of Technology. Results show that extrapolation methods may increase the RSS estimation error sometimes because it is very hard to predict the outside range. However, with more accurate extrapolation, the error would decrease. The performances of natural neighbor, linear and cubic interpolation are similar. The highest impact on RSS estimation comes from the extrapolation.

PREFACE

This Master of Science Thesis has been carried out in the Department of Electronics and Communications Engineering (ELT) at Tampere University of Technology (TUT), Tampere Finland.

I am pleased to express my sincere thanks and appreciation to my supervisor, Dr. Tech. Elena- Simona Lohan , who has given me the opportunity to study in WLAN-based positioning field and offered me valuable guidance. I would also like to express my appreciation to my other examiner, Prof Pertti Järventausta for the huge support. I am also grateful to Tianshu Chen, Jukka Talvitie and Wenbo Wang for their help. They really supported me a lot during this master thesis period.

Finally, I must say thanks to my parents and my friends for their love and support to complete my master degree.

Tampere, 4th Feb, 2015
Jindou Chen

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

ANC Adaptive Normalized Convolution

AOA Angle of Arrival

AP Access Point

API Application Programming Interface

BS Base Station

dBm Decibel Miliwatt

DOA Direction of Arrival

GPS Global Positioning System

IDW Inverse-Distance Weighted

MMPE Modified Minimal Polynomial Extrapolation

MPE Minimal Polynomial Extrapolation

RBF Radial Basis Function

RRE Reduced Rank Extrapolation

RSS Received Signal Strength

RSSI Received Signal Strength Indicator

TDOA Time Difference of Arrival

TEA Topological ε Algorithm

SEA Scalar ε Algorithm

TEA Topological ε Algorithm

TPS Thin Plate Spline

TUT Tampere University of Technology

VEA Vector ε Algorithm

WLAN Wireless Local Area Network

WiFi	Wireless Fidelity
P_T	Transmit power of Access point
Ψ	Shadowing deviation
n	Path-loss exponent
c	velocity of light in free space
λ	wavelength
f	frequency
$\|\|.\|\|$	Euclidean norm

1. INTRODUCTION

1.1 Background and motivation

In this information era, people require to get access to their exact location and companies are hard-pressed to figure out the exact customer location, for applications such as user-tailored advertising, mobile office tracking, and transport optimization. They also need the navigation service anywhere anytime.

Satellite positioning, such as the Global Positioning System (GPS), offers good location accuracy and availability in outdoor scenarios. Indeed, the satellite signals can be received with good power levels outdoors; however, in indoor environments, these signals are typically received at very low power, because of the walls and ceilings attenuations. The need for a better solution for indoor environment has led to the attention on WLAN-based positioning [1].

There are several advantages of WLAN-based positioning. Firstly, WLAN infrastructure is already existent. Moreover, the coverage is very high in many areas. Thirdly, WLAN positioning has entered the standardization already, for example, 802.11k and 802.11v standards. Finally, in WLAN, there are a plenty of emitters or Access Points (AP) and the Received Signal Strength (RSS) values are easily accessible from the Application Programming Interface (API).

Nevertheless, depending on the embraced positioning method, there are various errors in WLAN-based positioning. For example, during the data collection phase, different devices may give different values, thus calibration is typically needed. This source of error is seen as a measurement error. Similarly, when path loss models are used for the position estimate, the different path-loss models may have modeling errors and there is always random shadowing or random fluctuation of RSS. There are also floor detection errors and calibration errors in WLAN-based positioning. These errors are issues waited to be solved.

In this thesis, we focus on applying interpolation and extrapolation methods in WLAN training databases, in order to decrease the positioning errors. Generally, the interpolation and extrapolation methods are widely used in image processing field. However, this time they are used in the positioning field.

1.2 Thesis objectives

This thesis studies different performances of interpolation and extrapolation methods using RSS measurements. The objectives have been:

1. Understanding the basic RSS-based positioning methods, such as fingerprinting and path-loss models.
2. Study the popular interpolation and extrapolation methods in the context of WLAN RSS-based positioning. In the interpolation stage, we focused on triangulation-based methods.
3. Investigate the performance for each interpolation and extrapolation method with the measurement data collected in 3 different campus buildings and draw comparative conclusions about the best interpolation and extrapolation methods.

1.3 Author contributions

The major contributions of this thesis are:

1. Literature review of RSS-based positioning methods, interpolation and extrapolation methods and their applications.

2. Implementation in Matlab of the files needed for analyzing the performance of different interpolation and extrapolation methods based on the data collected from several buildings in Tampere University of Technology.

3. Modification of the Matlab simulator initially provided by the TUT team, in order to decrease the measurement error.

4. Investigating and analysis of the performance of interpolation and extrapolation methods with measurement data.

1.4 Thesis organization

The rest of this thesis is organized as below:

Chapter 2 introduces different WLAN-based positioning methods. RSS-based fingerprinting and path-loss models are explained in details.

Interpolation and extrapolation methods are describes in Chapter 3 and 4. In chapter 3, we focused on triangulation-based interpolation methods.

Chapter 5 introduced some basic concepts about extending the two dimensional models to three dimensional interpolation and extrapolation models.

The measurement description and data analysis of the campus building are addressed in Chapter 6 and Chapter 7.

Conclusions and future works are discussed in the final Chapter 8.

2. PRINCIPLES OF WLAN-BASED POSITIONING METHODS

Different measurements taken from radio frequency signals between mobile station (MS) and the fixed stations, namely the Base Stations (BS) or Access Points (AP) are used on different wireless location systems. The measurements can be: angle-of-arrival (AOA), time-of-arrival (TOA), time-difference-of-arrival (TDOA), and received-signal-strength (RSS).

2.1 Angle-of-Arrival (AOA)

The Angle of Arrival (AOA) technique is also called Direction of Arrival (DOA). This technique estimates the position by using geometric relationships, such as triangulation method. The user position can be calculated by intersecting two lines passing through the BS or AP with the measured angles. In this method, non line of sight will cause problem. The technique is illustrated in Figure 2.1.

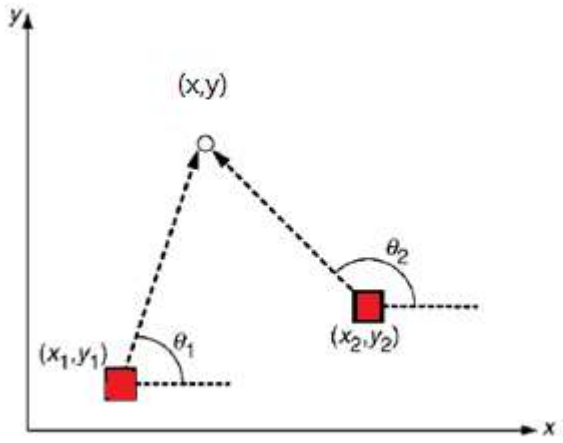


Figure 2.1 *Two base stations and user and angles measured to find user position*

The points (x_1, y_1), (x_2, y_2) are the coordinates of two APs or BSs, angles θ_1, θ_2 are the AOAs and they are assumed to be known. Equation 2.1 is the illustration of the relationship between the user position and base station positions [2].

$$\frac{y - y_i}{x - x_i} = \tan(\theta_i), \text{i=1,2} \tag{2.1}$$

By substituting two given points and angles, the value for (x, y) can be obtained.

2.2 Time-of-Arrival (TOA)

This method measures the travel time of a radio signal from a transmitter to a remote receiver to locate the position. In order to achieve this goal, at least 3 transmitters are needed. The location can be found at the intersection of the three circles shows in Figure 2.2 .

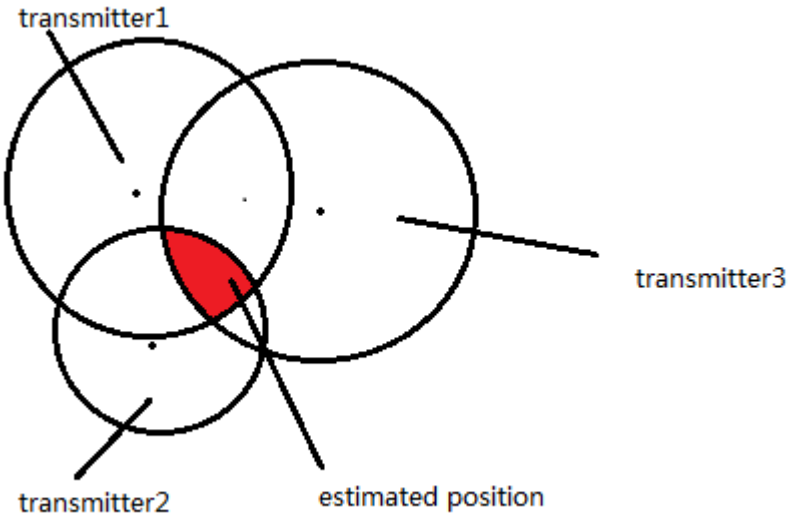


Figure2.2 *Time of Arrival*

The problem in this method is that, all the receivers and transmitters should be synchronized. The difference between arrival time and departure time is a significant factor in TOA method, and the error of difference due to clock offset should avoided. However, considering the high speed the signal travels, this error is hard to avoid.

2.3 Time-Difference-of-Arrival (TDOA)

Time-Difference-of-Arrival (TDOA) method is similar to TOA, but it uses the time differences instead of the absolute time. The travel time from transmitter to receiver is used to measure the distance. However, in TDOA, the travel time difference between transmitters is calculated to find the distance between those transmitters and the receivers. Therefore, hyperbolic positioning or multilateration are used in this method to locate the position. To make the measurement accurate, synchronicity is also required in TDOA.

2.4 Received Signal Strength Methods

RSS (Received Signal Strength) values represent the power of incoming signal in a receiver in decibel (dB) or decibel miliwatt (dBm). Generally, a stronger signal means a higher RSS value. Though different manufacturers have distinct design, each receiver

measures a Received Signal Strength Indicator (RSSI). This is then mapped to RSS value in dBm.

RSS value can be used in range-based algorithms. For example, in a trilateration method, RSS values can be used to acquire the distance between transmitter and receiver.

RSS value can also be used in database comparison algorithms, such as path-loss method and fingerprinting method. Unlike the trilateration method, no distance measurements are needed. Instead, the RSS-based localization method has two stages: training phase and estimation phase. Next section explains how these two stages applied in path-loss models and fingerprinting method separately.

2.4.1 RSS-based fingerprinting method

The comparison between the RSS of the receiver and the data measurements collected off-line is made firstly. Then the point which matches best in terms of RSS is figured as the location. This process is called fingerprinting. As it was mentioned above, there are main two stages [2] in fingerprinting method, namely the training phase and the estimation phase.

1. Training phase

A database is collected during this phase. Typically, the database is in the form of received signal strengths and measurements coordinates. The RSS is measured depending on the heard Access Points in the specific location. The coordinates of the fingerprints are measured according to the off-line maps made initially. In the training phase, the information of RSS and the coordinates of each fingerprints and each heard AP are stored.

Figure 2.3 shows the points heard from one AP in the first floor of A university building. x and y axis represents the *(x, y)* coordinates of the measurements. z axis represents the RSS value which the measurement heard from the AP. In Figure 2.3, these measurements are heard from the 20^{th} ap at the second floor, nf refers to the floor of that building and ap refers to the number of the ap.

2. Estimation phase

During this phase, the comparison between the RSS of user point and the RSS in database is made. According to the comparison, the point in database which is closest match with the user is determined. The point which has the smallest Euclidian distance [1] is returned as the estimated position. Sometimes, the result of the estimated position comes from single fingerprints. Sometimes, it comes from the average of several closest neighbors fingerprints.

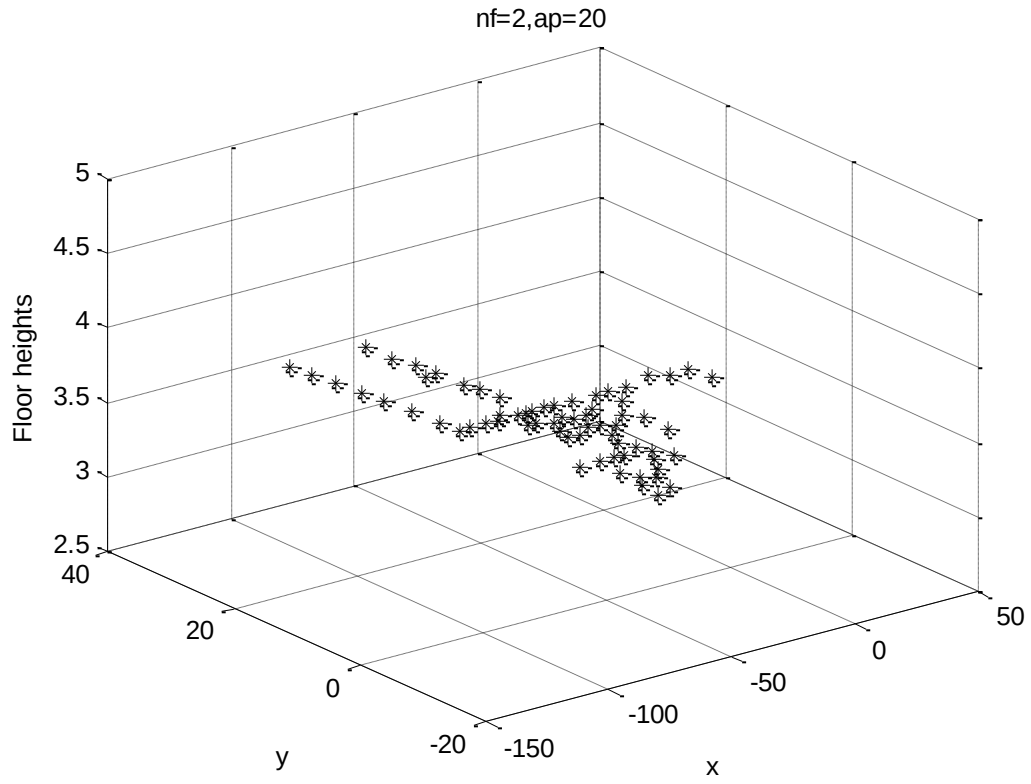


Figure 2.3 *Measurement points where one AP is heard in the first floor in A university building*

2.4.2 RSS-based path-loss models

Path loss means the difference between the measured RSS and the transmitted signal power. The path loss model is the method based on maximum likelihood estimation. It also has training phase and estimation phase.

1. Training phase

During the training phase, the distance between the transmitter and receiver is estimated, typically through a path-loss model. The model below shows an example of the relationship between RSS and distance. It includes at least path losses and shadowing.

The shadowing means the random fluctuations in signal strength.

- Free space path-loss models:

  This model is applied when there is visible LOS path between transmitter and receiver. The free space received power is modeled as:

$$P_R = P_T G_T G_R \left[\frac{\lambda}{4\pi d} \right]^2 \quad (2.2)$$

where, P_T is transmitted power, G_T is the transmitted antenna gain, G_R is the receiver antenna gain. λ is the wavelength of transmitted signal.

The free space path-loss model (in dB) is given as:

$$L = 10\log_{10}\left[\frac{4\pi d}{\lambda} \right]^2 = 32.4 + 20\log_{10} d_{km} + 20\log_{10} f_{MHz} \quad (2.3)$$

where , d_{km} is the distance between transmitter and receiver. f_{MHz} is the carrier frequency.

- Simplified path loss model

The received power is:

$$P_R(d) = P_R(d_0)\left(\frac{d_0}{d} \right)^n \quad (2.4)$$

where, d_0 is reference distance, $P_R(d)$ is the received power, n represents path-loss exponent.

Table 2.1 shows the path-loss exponent in different environments according to [1]. However, these are only some examples and the values can vary in various environments from the values reported here

Table2.1 *Some examples on path-loss exponent in different environments, according to [1]*

Environment	**n (Path-loss Exponent)**
Free space	2
In building LOS	1.6-1.8
Obstructed in building	4-6
Obstructed in factories	2-3
Urban area cellular Radio	2.7-3.5
Shadowed urban cellular radio	3-5

The path loss (in dB) is given as equation 2.5,

$$L = 10n \log_{10} \left\lceil \frac{d}{d_0} \right\rceil \tag{2.5}$$

- Floor and wall factor model

This model is meant for indoor propagation environment and it was introduced in [1]. This model characterizes indoor path losses by free space attenuation, additional loss factors related to a number of floors and walls intersected by the straight-line to the receiving terminal. The model is written in the form of equation 2.6:

$$L = L_{ref} + 20 \log_{10} d + \sum_{f=1}^{F} FAF(f) + \sum_{w=1}^{W} WAF(\mathrm{w}) \tag{2.6}$$

where ,

L_{ref} -Reference path loss at d=1 m distance

$FAF(f)$ -floor attenuation factor for floor f

$WAF(\mathrm{w})$ -wall attenuation factor for wall w

- Okumura-Hata model and COST 231 model

They are applied in outdoor environment. They are used for coverage prediction in macrocells. This empirical path-loss model is based on the measurement data made by Okumura in Tokyo in 1968. It is valid for frequencies between 150 MHz and 1500 MHz. Okumura-Hata model was extended by the European co-operative for scientific and technical research (EURO-COST) from 1500 MHz to 2GHz. This was called COST 231 model, the model is given by equation 2.8,

$$L = A + B \log_{10}(f_{MHz}) - 13.82 \log_{10}(h_b) + (c - 6.55 \log_{10}(h_b)) \log_{10}(d_{km}) - k \tag{2.7}$$

where,

h_b -base station antenna height [m]

d_{km} -link distance [km]

f_{MHz} -carrier frequency [MHz]

c- tunable parameter [44-47]

K -correction factor [default = 0]

The frequency dependent parameters A and B are given in Table 2.2:

Table2.2 *A and B parameters*

	150-1000MHz	1500-2000MHz
A	69.55	46.3
B	26.16	33.9

Corresponding to different models, during the training phase, path loss parameters such as transmit power, path loss coefficient correction factor, frequency dependent parameters, etc, are estimated.

2. Estimation phase

During the estimation phase, trilateration approach [30] can be used to calculate the user position.

Suppose the location of AP or BS is known as (x_{ap}, y_{ap}, z_{ap}), *(x,y,z)* denotes the MS location, then

$$d_{ap} = \sqrt{(x - x_{ap})^2 + (y - y_{ap})^2 + (z - z_{ap})^2} \tag{2.8}$$

where d_{ap} is the distance between MS and AP/BS.

If the AP/BS location (x_{ap}, y_{ap}, z_{ap}) is not known, it also can be estimated in the training phase. For example, the coordinates (x_{ap}, y_{ap}, z_{ap}) can be obtained by taking an average over the positions of the measurements with the highest RSSI [5].

Thus, there are three unknown parameters in Equation (2.8), to obtain the distance, at least three APs/BSs should be known. With three or more transmitters, the exact location can be estimated. Figure 2.4 shows the illustration of the trilateration principle with 3 APs.

In Figure 2.4, the small green, red, purple circles represent three APs respectively. The black block is the location of unknown MS. The blue points are the measurements heard by APs.

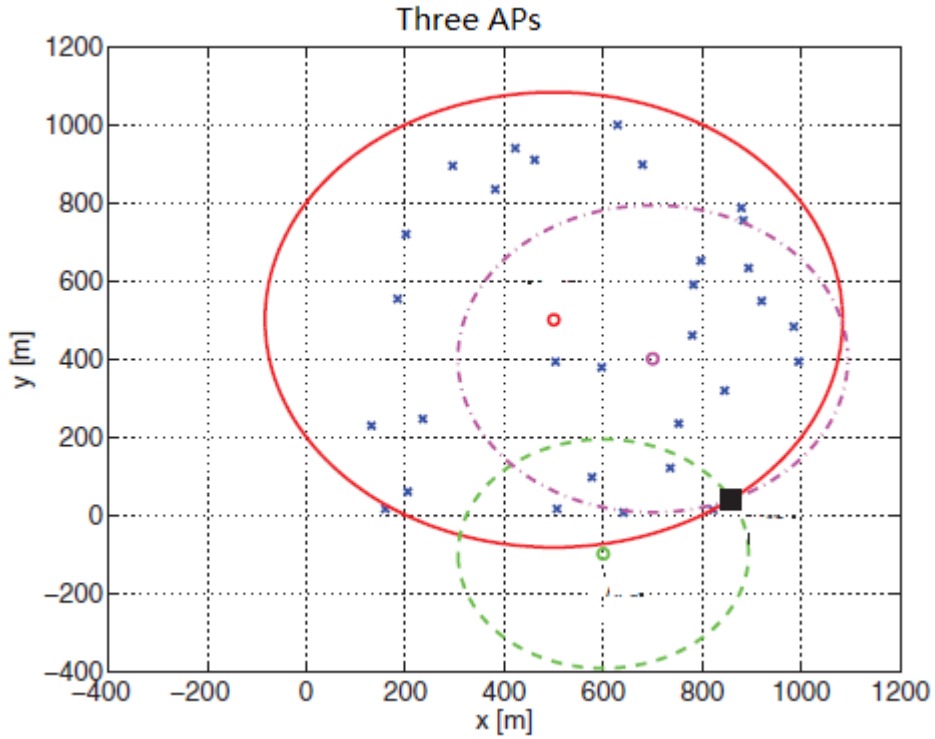


Figure 2.4 *Trilateration principle with 3 APs*

The major difficulty in trilateration approach is to obtain the real distance between SP/BS and MS. Distance, losses through walls and floors and other interference all lead to signal attenuation. Thus, it is not easy to obtain the distance accurately.

2.4.3 Path-loss model versus fingerprinting

Both methods collect the initial RSS data base during training phase. For fingerprinting method, the coordinated of the fingerprints and their RSS values are stored. In other words, huge database are required in this method. This is why fingerprinting is not suitable to global positioning. For path loss models, the size of the database is proportional to the number of parameters. It is available in large scales. More comparison between path loss model and fingerprinting are listed in Table 2.3.

Table 2.3 *Comparison between path-loss model and fingerprinting*

	Path-loss model	Fingerprinting
Data storage needs	Relatively low	Relatively high
Accuracy	Relatively low	Relatively high
limitation	Feasible for large areas	In feasible for large areas (requiring large amount of data transfer from/to mobile)

3. INTERPOLATION METHODS

In general, there are two categories of interpolation methods: global and local methods. In global interpolations, the computational cost increases with the amount of known data, because all the know data has to be used in this method. When there is new data added, a recalculation is required. In local interpolation, only the data which fall within a specific point's neighborhood are used to calculate output values. Therefore, when the sample numbers increase, the computational cost of this method does not grow quickly.

There are several interpolation methods listed, as below:

1) Triangulation based (e.g., natural neighbor, nearest neighbor, linear, and cubic)[3], [6], [7], [9]

2) Radial basis function (RBF) [19]

3) Spline interpolation [9]

4) Inverse-distance weighted (IDW) [8]

6) Adaptive normalized convolution (ANC) [3]

In this chapter, after introducing some basic concepts, the triangulation based methods: natural neighbor, nearest neighbor, linear, and cubic interpolation are discussed in more details because these four methods are applied in the Matlab simulator in this thesis.

3.1 Voronoi diagrams and Delaunay triangulation

To get connectivity from scattered data, two concepts need to be introduced, namely Voronoi diagrams and Delaunay Triangulation. Voronoi diagrams and Delaunay triangulation in 2D space are discussed in this section. Their applications in 3D space are introduced in Chapter 5.

Figure 3.1 shows the Voronoi diagram in 2D space. Those blue Points from x_1 to x_{14} are known already. Based on these points, Voronoi cells V(x_i) are generated. The red point P is one point in Voronoi cell V(x_5).

We can observe that point P is closer to point x_5 than to any other points in Figure 3.1. This fact also true for any other point within V(x_5) . This is the principle for Voronoi diagram.

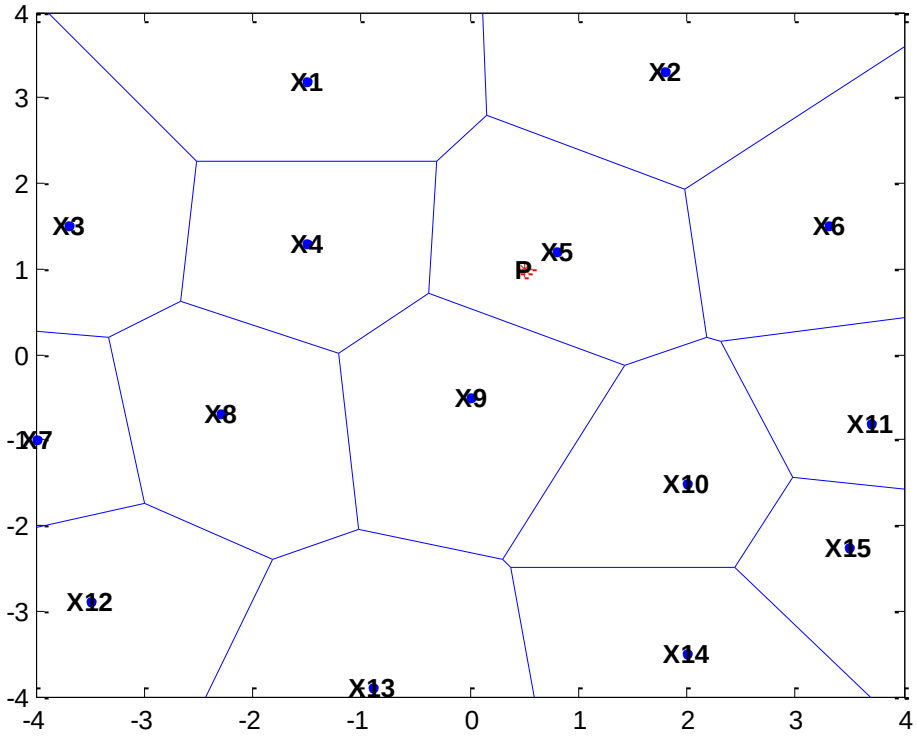


Figure 3.1 *Voronoi diagram in 2D*

Two cells share the same edge are called adjacent cells. For example, in Figure 3.1, V(x_4) and V(x_5) are adjacent cells because they share the same edge.

Delaunay triangulation is generated by connecting sites with those in adjacent cells. Those points are the vertices of triangulation. Figure 3.2 shows the Delaunay triangulation in 2D space.

A triangulation of scattered data points consists of vertices, edges and faces. Edges are the connection of two vertices. Faces are the connection of three vertices.

Given a set of points $X = \{x_{1,\cdots,} x_n\}$. There are three features for the 2D Delaunay Triangulation [3]:

1) Three points x_i, x_j, x_k in X belong to the face from Delaunay triangulation X if and only if no other points lies in the circle around x_i, x_j, x_k

2) Two points x_i, x_j from an edge if and only if the circle around x_i, x_j does not include any other points from X.

3) For each triangle the circumcircle does not contain another sample.

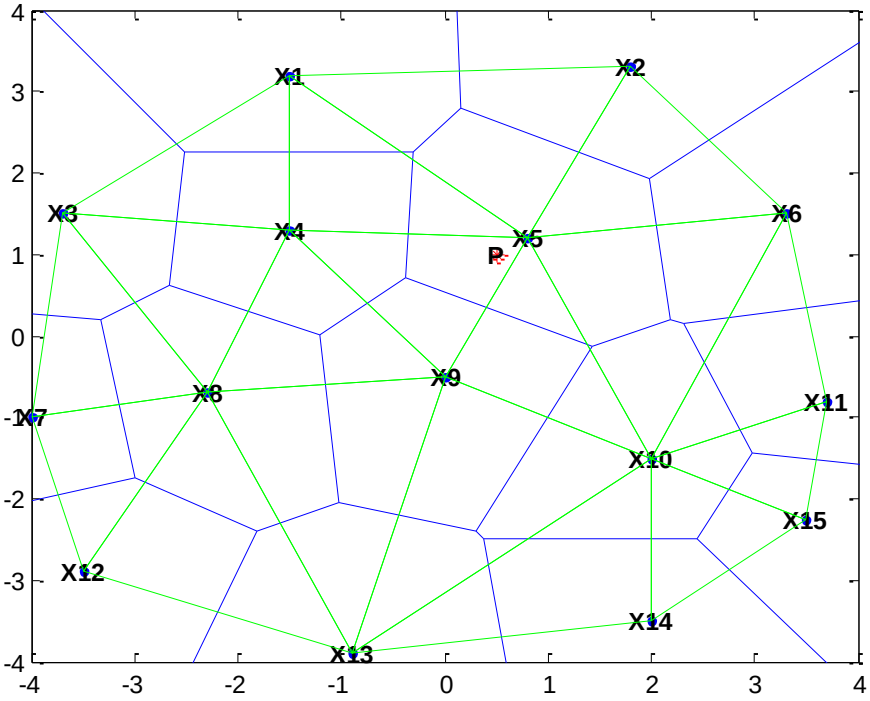
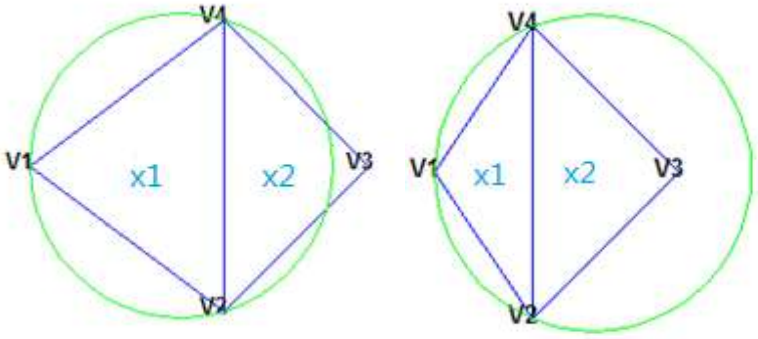


Figure 3.2 *Delaunay triangulation in 2D*

The features can be explained in Figure 3.3.



(a) (b)

Figure 3.3 *Delaunay triangulation features explanation*

V1, V2, V3, V4 are the vertices of the Delaunay triangulation. $^{\Delta V1V2V4}$ is triangulation X1. $^{\Delta V1V2V3}$ is triangle X2. In (a), the circumcircle associated with X1 is empty. It does not contain any point in its interior. However, in (b), the circumcircle associated with X1 is not empty. It contains V3 in its interior.

3.2 Triangulation based Interpolation methods

In this chapter, four triangulation-based methods are presented: nearest neighbor, linear, natural neighbor and cubic interpolation methods. The general applications of these methods are showed in Table 3.1as below.

Table 3.1 *Genera application for triangulation based interpolation methods*

Methods	Description
Nearest neighbor	Supporting 2-D and 3-D
linear	Supporting 2-D and 3-D
Natural neighbor	Supporting 2-D and 3-D (efficient trade of between linear and cubic)
cubic	Supporting 2-D only

3.2.1 Natural neighbor interpolation method

Natural neighbors are defined as the Voronoi cells from two sites which share a common edge. The natural neighbor interpolation method is applied in two or three or even higher dimensions. This method was proposed by Sibson in 1981[3]. It was used for data smoothing and approximation.

There are different form of natural neighbor interpolant according to different weighted averaging, such as Sibson's interpolant and Laplace interpolant and so forth. Among those methods, Sibson's interpolant is the most popular one [3].

Suppose a set of points $X = \{x_{1,\cdots,} x_n\}$ are defined in 2D space. In WLAN context, the coordinates (x, y) of point x_i refers to the x axis and y axis value of each measurement. $f(x_i)$ refers to the power in which this measurement was heard by the particular AP. The power was represents by RSS values in dB.

The Sibson's interpolant in natural neighbor interpolation is defined as below:

$$f(x) = \frac{\sum_i a_i \cdot f(x_i)}{\sum_i a_i} \tag{3.1}$$

In this formula, a_i represent the overlap area between Voronoi cells $V(x_i)$ and $V(x)$.

This method is illustrated in Figure 3.4. In Figure 3.4, Voronoi cells and $V(x_1), V(x_2), V(x_3), V(x_4), V(x_5)$ are adjacent cells, because they share the same edge. Points x_1, x_2, x_3, x_4, x_5 are the natural neighbors. The blue area is the overlapping area

corresponding to x_1 , which represents a_i in formula (3.1).

The interpolation value at x is determined by the weighted averaging of those natural neighbors.

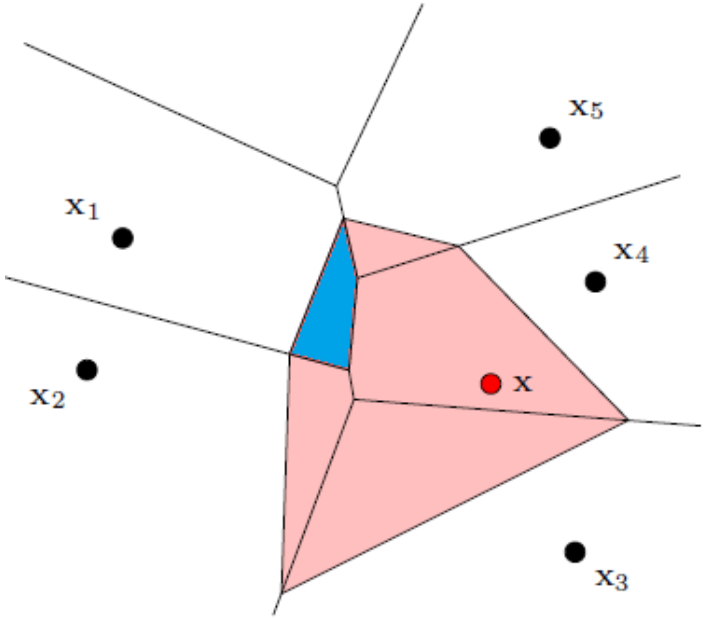


Figure 3.4 *Natural neighbor interpolation*

Generally, based on Voronoi diagram and Delaunay triangles, in order to interpolate a value $f(x_i)$ at point x, three steps should be followed. Firstly, we have to determine the location of the point x. In order words, it means to find the Delaunay triangulation which contains this point. Secondly, we have to determine the triangle which is not a Delaunay triangle after the point x added. Thirdly, Sibson's interpolant is obtained based on formula (3.1).

Figure 3.5 shows the measurement points heard by a specific AP at fourth floor in A university building in TUT. The red circles showed the coordinated of each AP with their RSS values. nf=4 refers to the fourth floor of that building and ap=2 refers to the 2th ap from that floor. Figure 3.6 shows the power maps after natural neighbor interpolation based on the measurements showed in Figure 3.5. Different colors represent different RSS values.

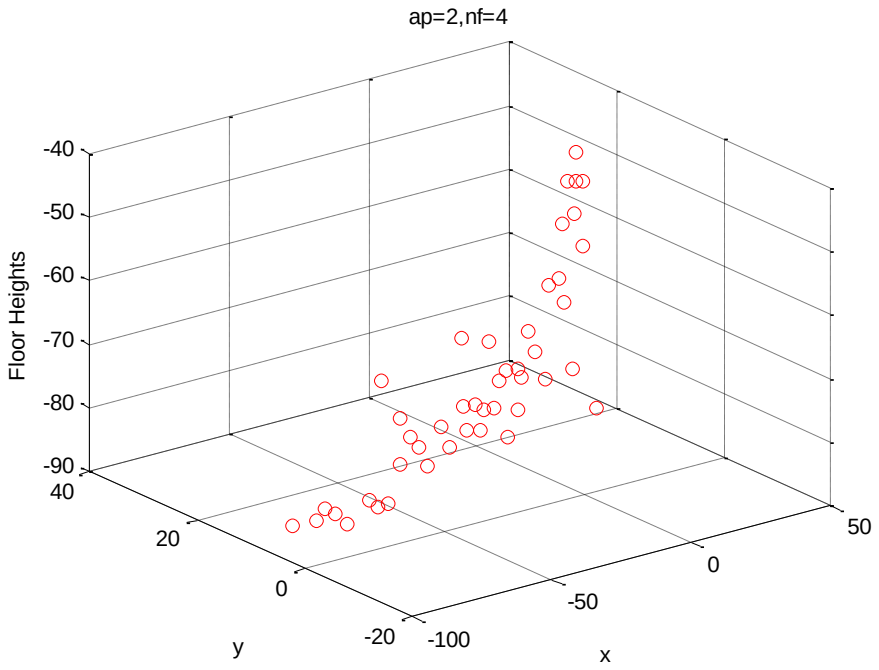


Figure 3.5 *Measurements points in which one AP is heard at fourth floor in A university building*

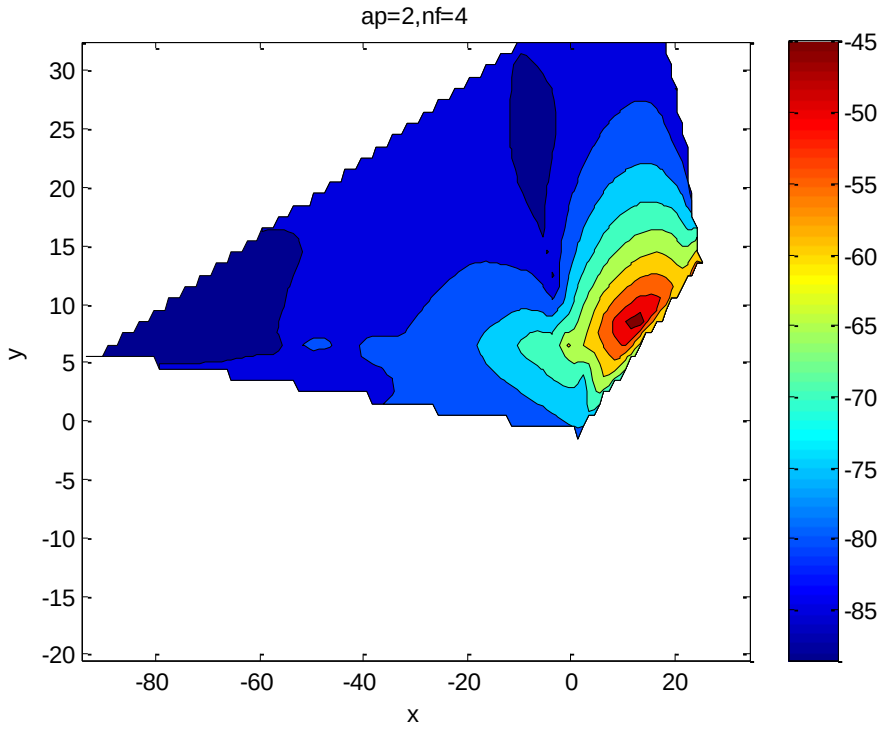


Figure 3.6 *Power maps after the natural neighbor interpolation based on the measurements of Figure 3.5*

3.2.2 Nearest neighbor interpolation

Compared to other interpolation methods, nearest neighbor interpolation has rather a simple implementation.

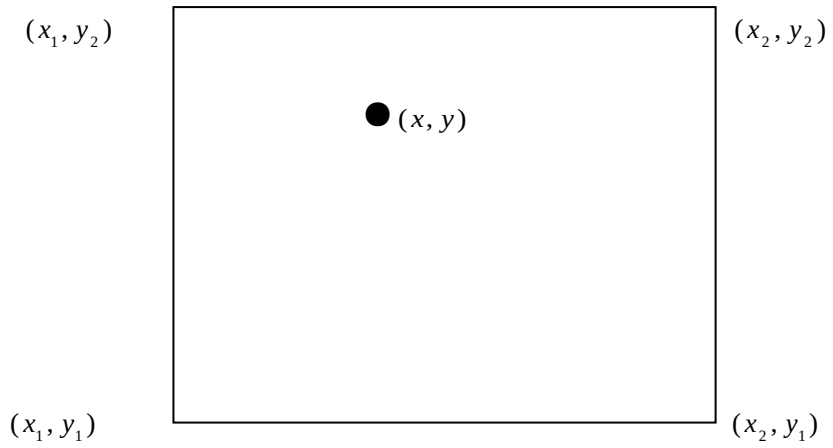


Figure 3.7 *Principle of Nearest neighbor interpolation*

As illustrated in Figure 3.7, here (x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2) are the 4 neighbor points to (x, y), and their values are $f(x_1, y_1)$, $f(x_1, y_2)$, $f(x_2, y_1)$, $f(x_2, y_2)$. The distance between (x, y) and (x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2) were calculated, then the values of (x, y) was set as the value of the point which is nearest to (x, y).

In Figure 3.8, point P is the interpolated point. A is one Voronoi cell in the whole diagram. All the points inside Voronoi cell A should be considered when calculating the interpolated point [7]. The value of the nearest point is saved for the interpolant.

Figure 3.9 shows the power maps after nearest neighbor interpolation based on the measurements showed in Figure 3.5.

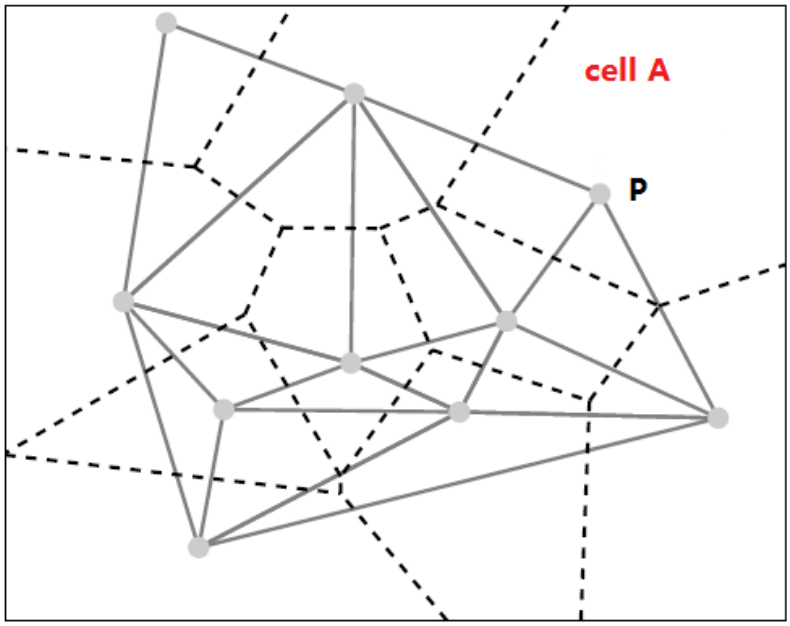


Figure 3.8 *Nearest neighbor interpolation in Voronoi diagram*

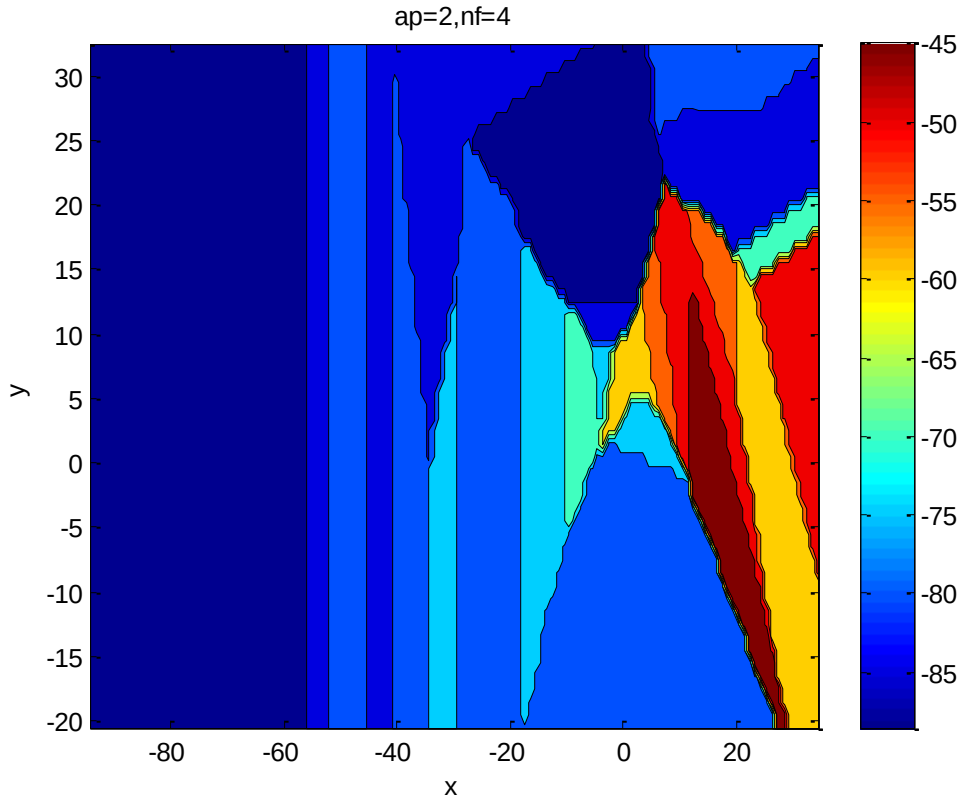


Figure 3.9 *Power maps after Nearest neighbor interpolation based on the measurements of Figure 3.5*

3.2.3 Linear interpolation

Compared to the Nearest neighbor method, the linear interpolation is only slightly more complex.

In one dimension space, only two points are needed to calculate the interpolated point value, as illustrated in Figure 3.10.

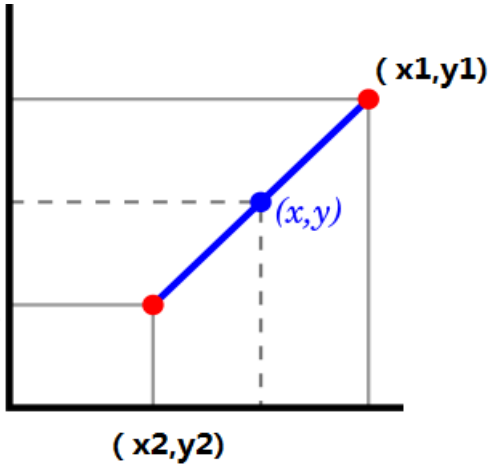


Figure 3.10 *Linear interpolation in one dimension*

In this Figure, two red points (x_1, y_1) , (x_2, y_2) are given. The point*(x, y)* is the interpolated point, and its coordinates are calculated via:

$$y = y_2 + (y_2 - y_1)\frac{x - x_1}{x_2 - x_1} \tag{3.2}$$

By substituting the value of two known points into this formula, the value for interpolated point can be obtained.

In 2-D spaces, four surrounding points are necessary to calculate the interpolated value. The general idea is applying linear interpolation in each direction separately. As showed in Figure 3.11.

In Figure 3.11, four red points P21 (x_1, y_1) , P22 (x_1, y_2) , P11 (x_2, y_1) , and P12 (x_2, y_2) are given, the two blue points Q2, Q1 are obtained from red points by linear interpolation. The green point is one to interpolate eventually.

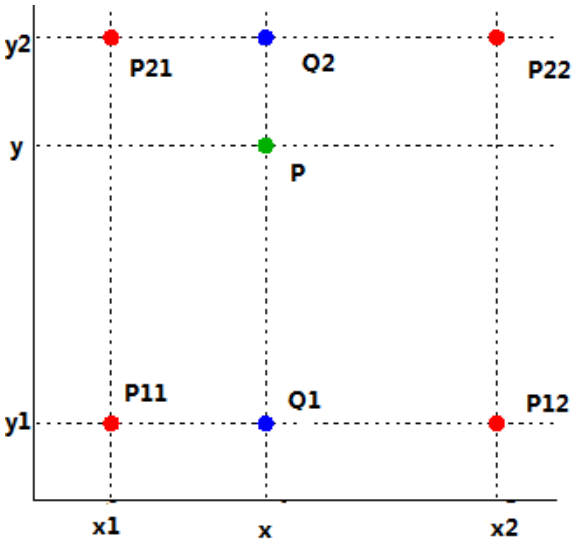


Figure 3.11 *Linear interpolation in two dimension*

Firstly, the linear interpolation was done at x direction,

$$f(Q_1) \approx \frac{x_2 - x}{x_2 - x_1} f(P_{11}) + \frac{x - x_1}{x_2 - x_1} f(P_{21}) \tag{3.3}$$

$$f(Q_2) \approx \frac{x_2 - x}{x_2 - x_1} f(P_{12}) + \frac{x - x_1}{x_2 - x_1} f(P_{22}) \tag{3.4}$$

Next, the linear interpolation in y direction was determined,

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(Q_1) + \frac{y - y_1}{y_2 - y_1} f(Q_2)$$

(3.5)

Eventually, the interpolated point was calculated as below,

$$f(x,y) \approx \frac{f(P_{11})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y_2 - y) + \frac{f(P_{21})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y_2 - y) + \frac{f(P_{12})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y - y_1) + \frac{f(P_{22})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y - y_1)$$

$$= \frac{1}{(x_2 - x_1)(y_2 - y_1)}\big(f(P_{11})(x_2 - x)(y_2 - y) + f(P_{21})(x - x_1)(y_2 - y) + f(P_{12})(x_2 - x)(y - y_1) + f(P_{22})(x - x_1)(y - y_1)\big) \tag{3.6}$$

There is no difference if y direction calculated first.

Figure 3.12 explains the implementation in a triangle.

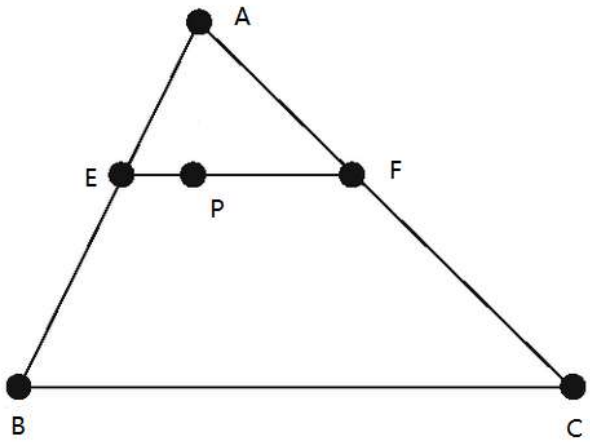


Figure 3.12 *Linear interpolation in triangulation*

In this figure, three vertices A, B, C are given. P is the interpolated point. The linear interpolation is first implemented at edge AB to obtain E. Next the same is done with the edge AC to obtain F. Finally, with the linear interpolation again, we have interpolated point P eventually.

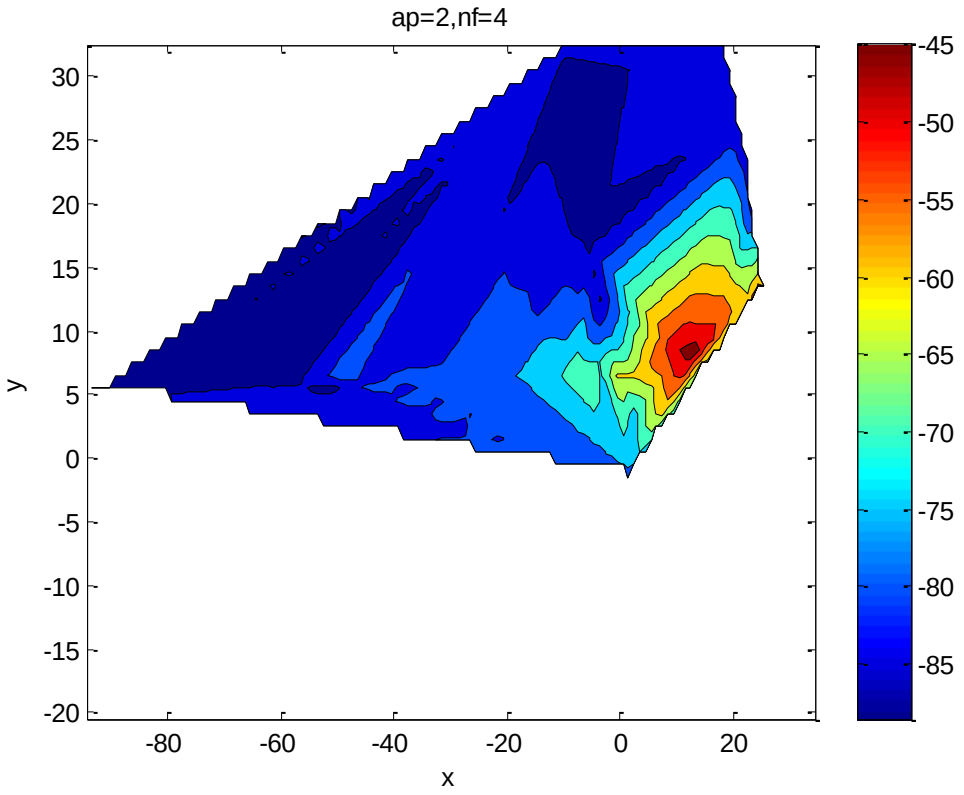


Figure 3.13 *Power maps after linear interpolation based on the measurements of Figure 3.5*

Figure 3.13 shows the power maps after nearest neighbor interpolation based on the measurements showed in Figure 3.5.

3.2.4 Cubic interpolation

Firstly, one dimension case in cubic interpolation is discussed. At least 2 points and their derivatives are required to obtain the model for this method.

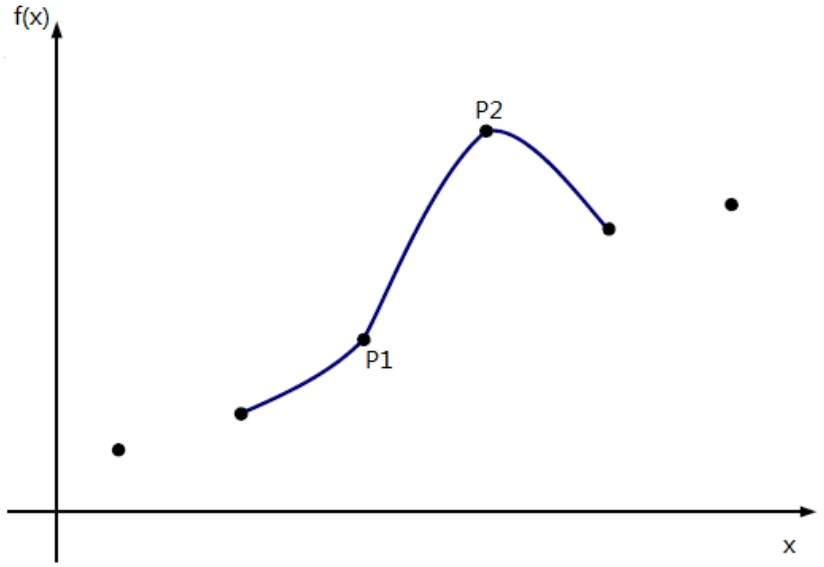


Figure 3.14 *Cubic interpolation in one dimension*

In Figure 3.14, points P1, P2 are given.

The models of cubic interpolation are:

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0 \tag{3.7}$$

$$\partial f(x) = 3a_3x^2 + 2a_2x + a_1 \tag{3.8}$$

Here $\partial f(x)$ refers to the derivative of $f(x)$, and a_3, a_2, a_1, a_0 are four unknown parameters, which need to be calculated.

Thus, in Figure 3.14 we have,

$$P(1) = a_3 x^3 + a_2 x^2 + a_1 x + a_0 \tag{3.9}$$

$$P2 = a_3 x_2^{\ 3} + a_2 x_2^{\ 2} + a_1 x_2 + a_0 \tag{3.10}$$

$$\partial P1 = 3a_3 x_1^{\ 2} + 2a_2 x_1 + a_1 \tag{3.11}$$

$$\partial P2 = 3a_3 x_2^{\ 2} + 2a_2 x_2 + a_1 \tag{3.12}$$

There are four equations, thus, parameters $a_{0,} a_{1,} a_{2,} a_{3,}$ can be solved by these four equations. Therefore, the model for this method is obtained.

Figure 3.15 shows the 2D cubic interpolation. P1, P2, P3, P4 are given. Their values are $(x_1, y_1), (x_2, y_2)$, $(x_3, y_3), (x_4, y_4)$ respectively.

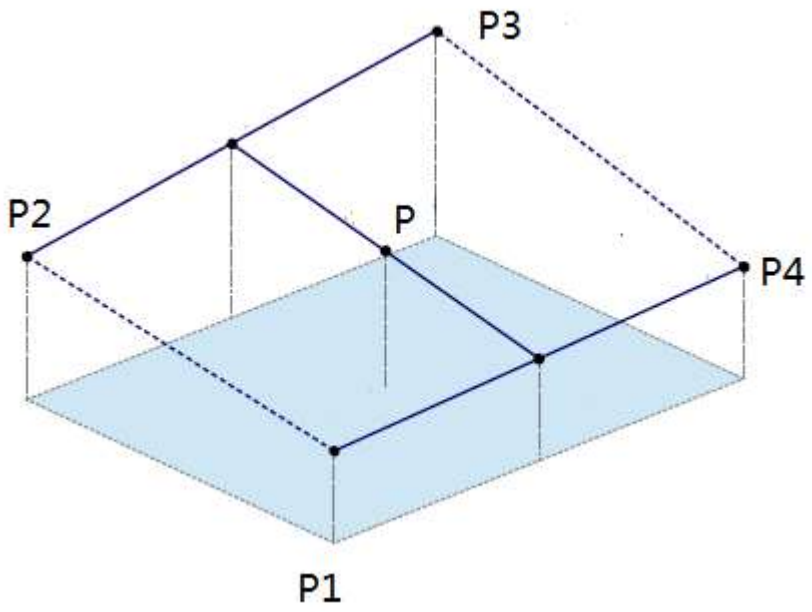


Figure 3.15 *2D cubic interpolation*

The model used in 2D cubic interpolation is listed as below [33],

$$f(x, y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j \tag{3.13}$$

$$\partial_x f(x, y) = \sum_{i=1}^{3} \sum_{j=0}^{3} i a_{ij} x^{i-1} y^j \tag{3.14}$$

$$\partial_y f(x,y) = \sum_{i=0}^{3}\sum_{j=1}^{3} j a_{ij} x^i y^{j-1} \tag{3.15}$$

$$\partial_{xy} f(x,y) = \sum_{i=1}^{3}\sum_{j=1}^{3} ij a_{ij} x^{i-1} y^{j-1} \tag{3.16}$$

In this model, $a_{ij}(i=1,2,3,4,\ j=1,2,3,4)$ refers to 16 known parameters need to be obtained. $f(x,y)$ in formula (3.13) represents the interpolated surface. Based on this formula, 4 equations are generated by 4 given points. $\partial_x f(x,y)$ represent the derivatives in x direction, $\partial_y f(x,y)$ represent the derivatives in y direction, $\partial_{xy} f(x,y)$ represent the cross derivatives. Thus, in total, 16 equations are required to solve the parameters for these models.

Figure 3.16 shows the power maps after cubic interpolation based on the measurements showed in Figure 3.5.

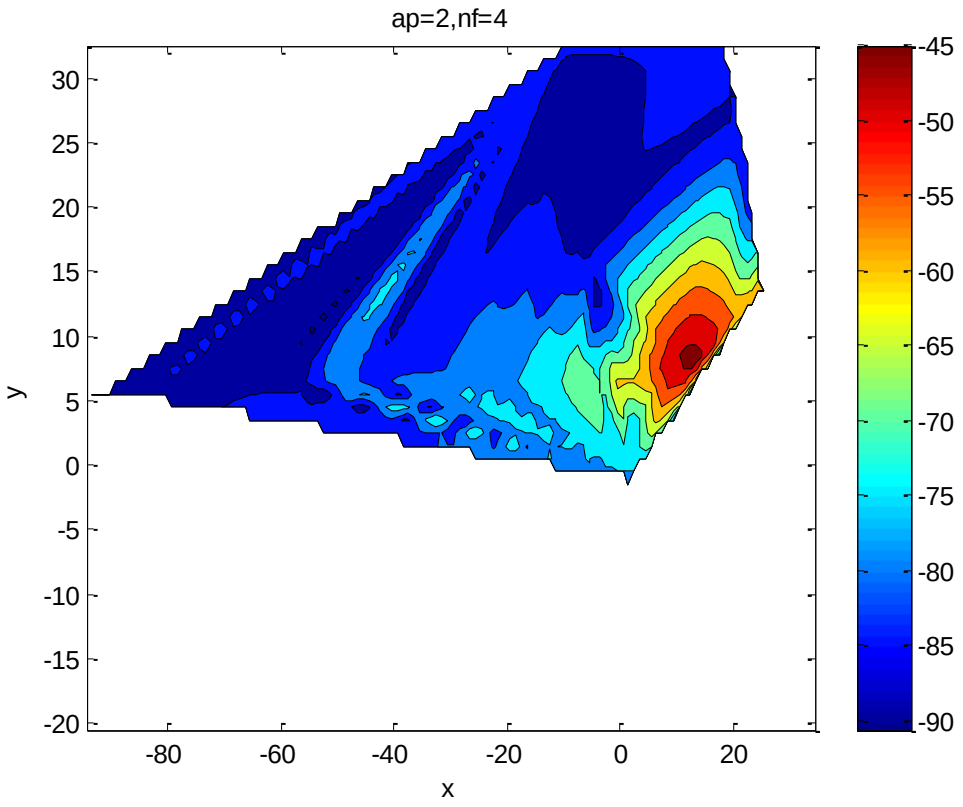


Figure 3.16 *Power maps after cubic interpolation based on the measurements of Figure 3.5*

3.3 Applications based on interpolation methods

Wireless Local Area Network (WLAN) gives the possibility of the connection between access point and wider internet. The development of wireless network is quick. Therefore, a better propagation model is required to be implemented in the networks. For example, in paper [14], linear interpolation helps to generate an advanced power signal received model in the campus environment. In the end, [14] gives the error between the propagation model and the actual values are 4.48% in dB. Because the propagation models applied in paper [14] are approximations of actual values, the author believed that this error was reasonable.

In the paper [34], comparison was made between various commonly used triangulation based interpolation methods for scattered-data interpolation over a range of sparcities, such as natural neighbor, nearest neighbor, linear and cubic interpolation. Other methods such as Adaptive Normalized Convolution (ANC), RBF and kriging are also tested in the paper. The results in [34] showed that natural neighbor and ANC have best performance in terms of giving the least error among the considered approaches, such as RBF, kriging, nearest neighbor, linear and the cubic interpolation and so forth. Among these methods, nearest neighbor and linear interpolation have worst performance due to bigger error they generated. The error was estimated between the points reconstructed by interpolation and the original points. However, it also mentioned that natural-neighbor interpolation is complex to implement. This result was similar with the result without extrapolation in this thesis. In the condition without extrapolation, the results of this thesis showed that natural neighbor has best performance while nearest neighbor has the worst.

4. EXTRAPOLATION METHODS

As we have discussed interpolation in chapter 3, interpolation is defined within the range of the data set. If the input is outside the range, this is extrapolation. Both interpolation and extrapolation are applied in the simulation in chapter 7.

Basically, extrapolation is harder than interpolation because of the higher uncertainty of the prediction.

In many engineering or other scientific applications, the accurate approximate limits of infinite sequence are required. However, in most conditions, the approximate limits converge slowly. Instead of computing a great number of terms to solve this problem, applying reasonable extrapolation methods to relatively smaller terms to achieve the high accuracy is a wise choice.

In this chapter, several 2D extrapolation methods are introduced.

4.1 Filling constant values

This method is rather simple. Depend on different conditions, a reasonable constant is chosen to fill the unknown points. The constant values can be zeros, means, minimum value or maximum value of the data set and so forth. This method is illustrated in Figure 4.2 and Figure 4.3.

Figure 4.1 shows the measurements heard by a specific AP at second floor in A university building.

Figure4.2 shows the power map of one floor in a campus building without extrapolation.

Figure4.3 shows the same floor power maps with extrapolation. In this power map, those points outside the range are filled with the minimum value of each Access point. Here different colors represent different RSS values.

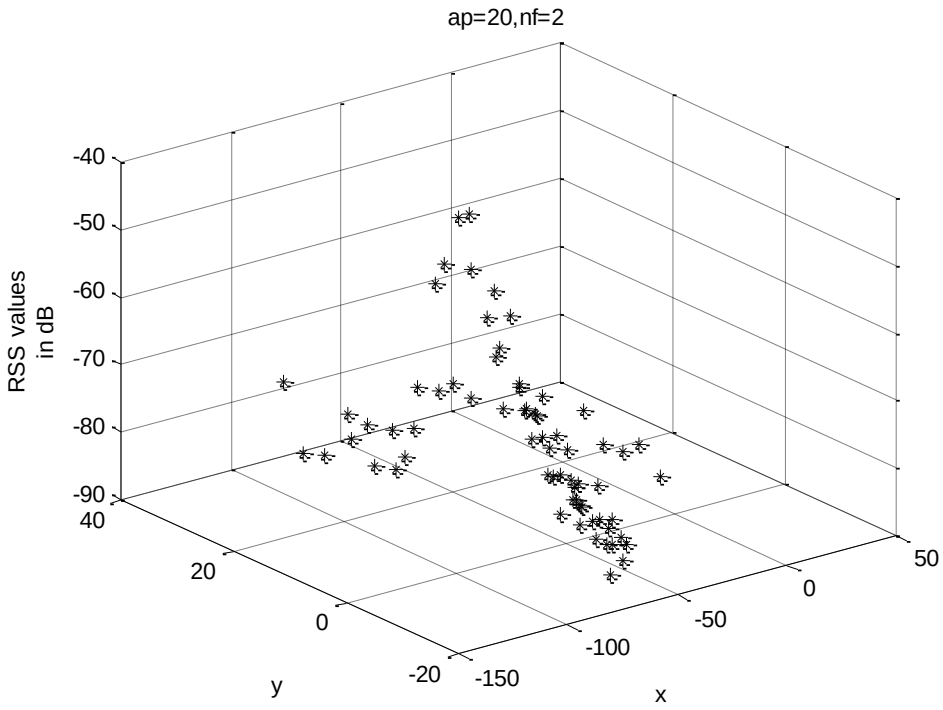


Figure 4.1 *Measurements heard by a specific AP at second floor in a university building*

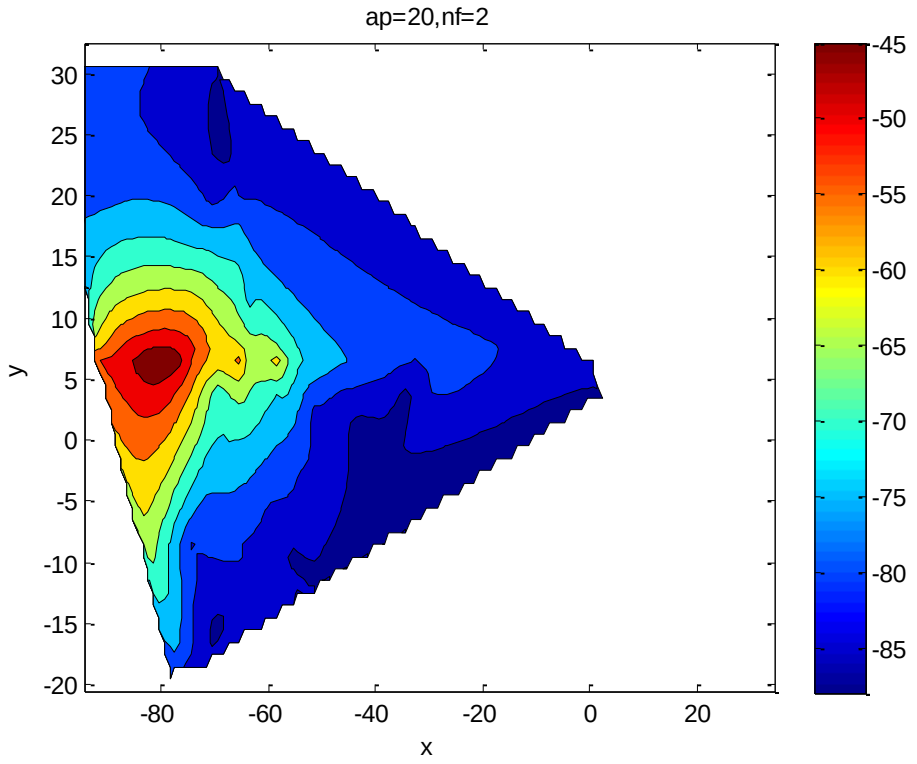


Figure 4.2 *Power maps without extrapolation based on the measurements heard by a specific AP at second floor in a university building*

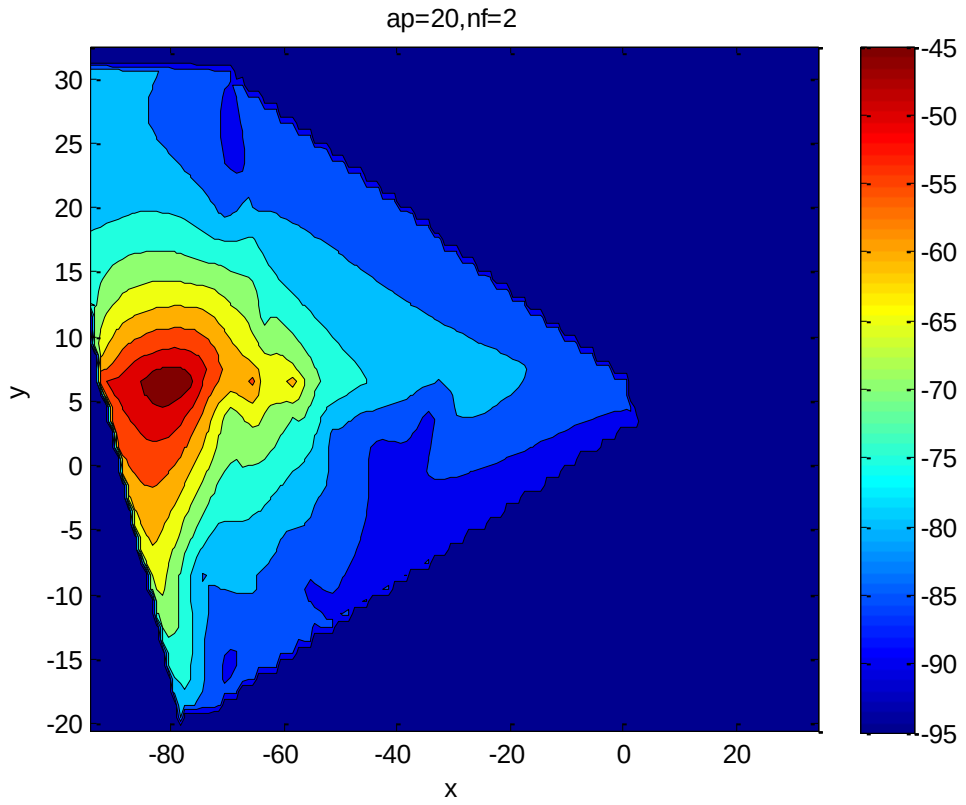


Figure 4.3 *Power maps with extrapolation (fill constant value) based on the measurements heard by a specific AP at second floor in a university building*

4.2 Linear extrapolation

The concept of this method is similar with the linear interpolation method. However, this method requires a linear function to predict values beyond the known data set.

Figure 4.4shows the power maps with linear extrapolation of the same floor with Figure 4.2.

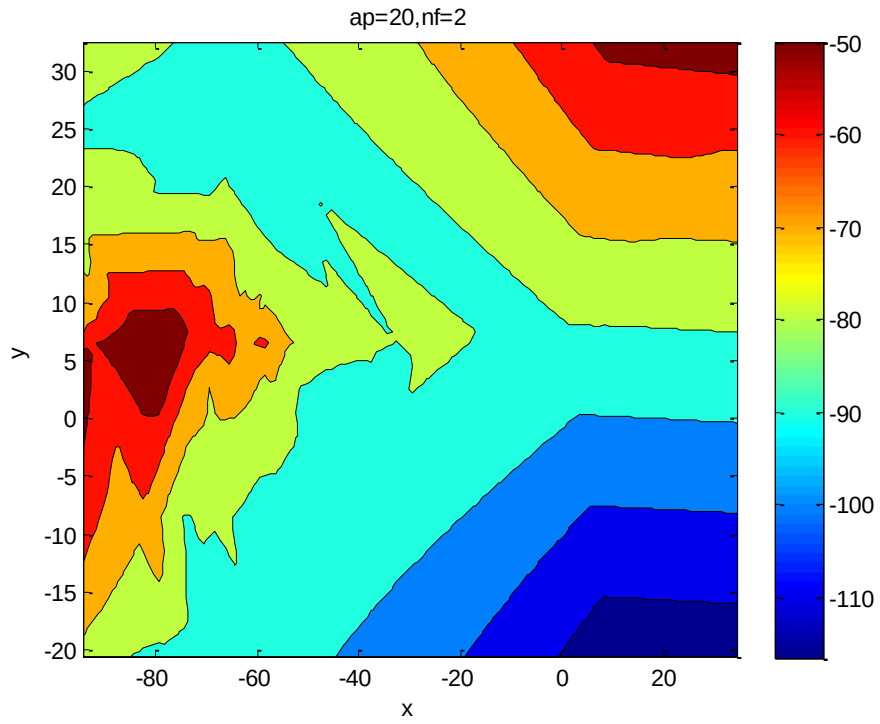


Figure 4.3 *Power maps with linear extrapolation based on the measurements heard by a specific AP at second floor in A university building*

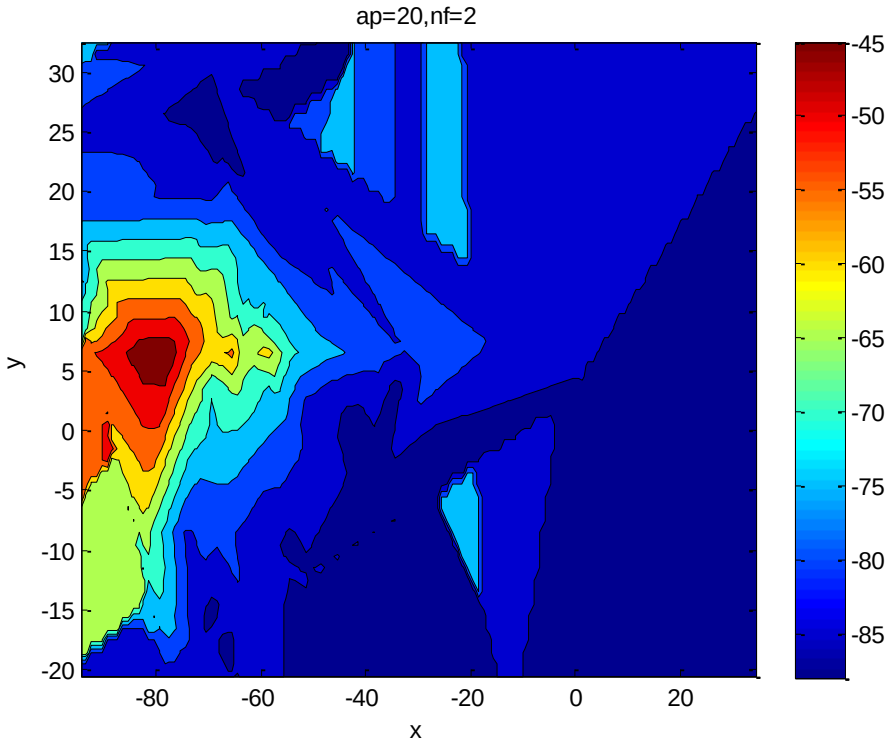


Figure 4.4 *Power maps with nearest neighbor extrapolation based on the measurements heard by a specific AP at second floor in A university building*

4.3 Nearest neighbor extrapolation

The concept of this method is similar with the nearest neighbor interpolation method. It would predict values beyond the known data set.

Figure 4.4shows the power maps with linear extrapolation of the same floor with Figure 4.1.

4.4 Vector Extrapolation

The principle of vector extrapolation is to transform a sequence of vectors to a new form which converge faster than the initial one.

Generally, there are two categories in vector extrapolation: the polynomial methods and the ε algorithms. In the polynomial methods, there are minimal polynomial extrapolation (MPE) method, the reduced rank extrapolation (RRE), the modified minimal polynomial extrapolation (MMPE) and so forth. In the ε algorithms class, there are topological ε algorithm (TEA) and scalar and vector ε algorithms (SEA and VEA) [11], [12].

Among those methods, the minimal polynomial extrapolation and the reduced rank extrapolation method are very efficient in accelerating the convergence of vector sequence. The principle of Vector extrapolation is really complex, it can be found in [11], [12].

There are also many other extrapolation methods are applied in the research field. For example, [10] proposed fast data-extrapolation methods to locate the narrow band in data extrapolation. Moreover, in [13], Gaussian extrapolation was used to approximate the unknown spot values in a specific region. This method normally applied when the space is not large. When the parameter space is relatively large, New-Raphson algorithm would be more effective. The principle of Gaussian extrapolation can be found in [13].

5. INTERPOLATION AND EXTRAPOLATION IN THREE DIMENSION SPACE

5.1 3D Delaunay Triangulation

Delaunay triangulation can partly avoid introducing distortions for interpolation. Also, this is kind of non-restrictive method while other methods may give unreasonable constraints on the shapes of the surface [32]. Due to the excellent adaptation to different geometric structures and data densities, Delaunay triangulation has advantages in 3D interpolation.

Compared to 2D, Delaunay triangulation in 3D has more complex structures, such as Delaunay tetrahedra and Voronoi polyhedral [20]. In other words, the computational cost is much higher in 3D. In Figure 5.1, it shows the 3D Delaunay triangulation generated by 20 random points.

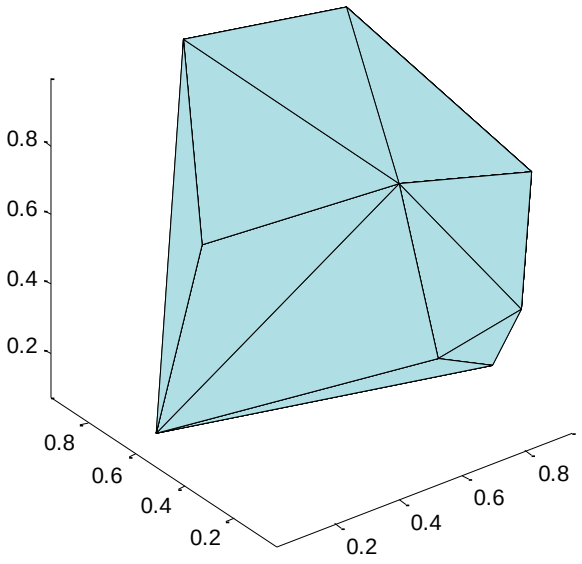


Figure 5.1 *3D Delaunay triangulation*

5.2 3D Linear interpolation

Linear interpolation in 3D is a straightforward extension from 2D linear interpolation. Figure 5.2 shows the linear interpolation in 3D space.

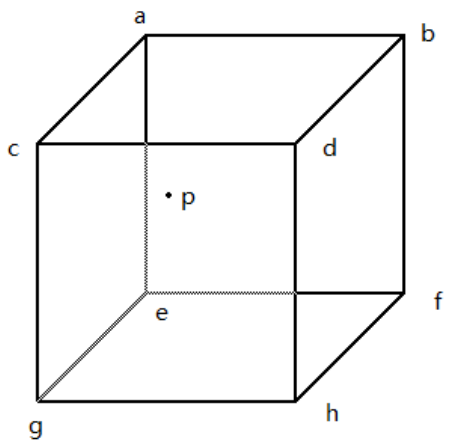


Figure 5.2 *3D linear interpolation*

In Figure 5.2, 8 corner vertices a, b, c, d, e, f, g, h are known. The coordinates of the interpolated point p is (α, β, γ).

The value for the interpolant can be calculated as:

$$f(\alpha, \beta, \gamma) = \mathrm{a} + \alpha(\mathrm{b} + \beta(\mathrm{e} + \mathrm{h}\,\gamma)) + \beta(\mathrm{c} + \mathrm{f}\,\gamma) + \gamma(\mathrm{d} + \mathrm{g}\,\alpha) \tag{5.1}$$

where the coordinates of a, b, c, d, e, f, g are (x_1, y_1, z_1) , (x_2, y_2, z_2) , (x_3, y_3, z_3) , (x_4, y_4, z_4) , (x_5, y_5, z_5) , (x_6, y_6, z_6) , (x_7, y_7, z_7) , (x_8, y_8, z_8) respectively.

Figure 5.3 shows a set of grid points. Figure 5.4 shows the same grid points after 3D linear interpolation.

Obviously, after interpolation, more points are added and the smoothness of the shape of Figure 5.4 is better than Figure 5.3.

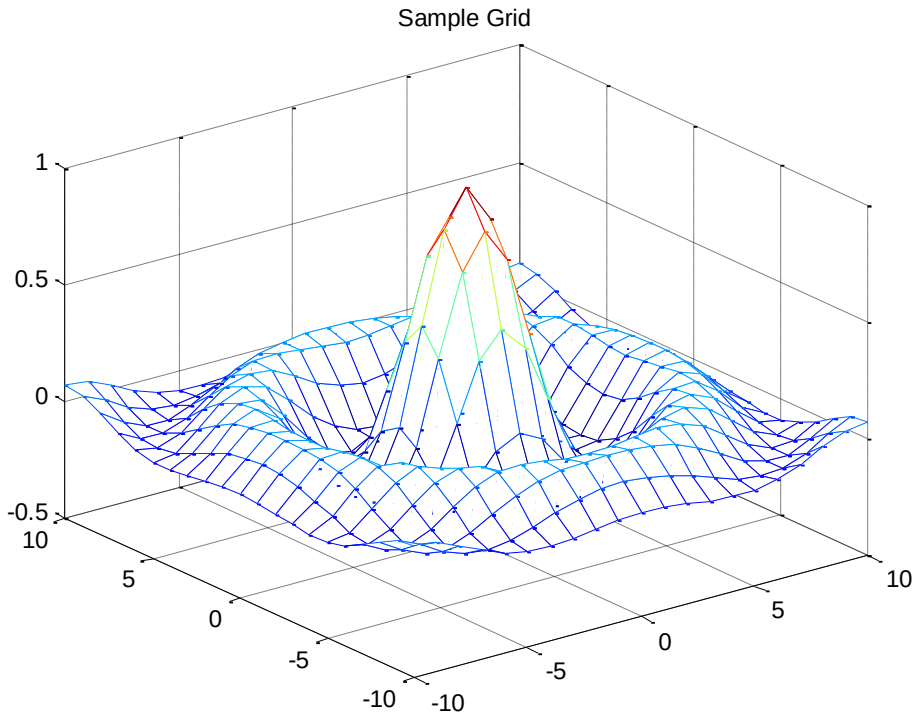


Figure 5.3 *Sample grid before interpolation*

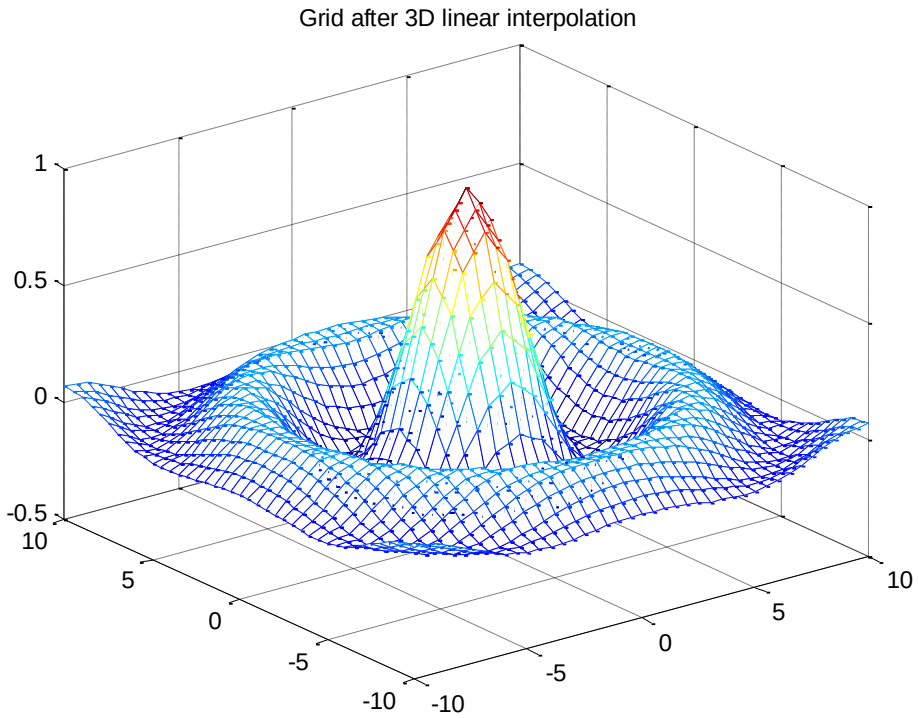


Figure 5.4 *3D linear interpolation based on sample grid*

5.3 3D cubic interpolation

The model of 3D cubic interpolation is similar to 2D cubic interpolation. In 2D model, there are 16 unknown parameters needed to be obtained. However, in 3D, the number of unknown parameters in the model increased to 64 as showed in Equation (5.2).

$$f(x,y,z)=\sum_{i=0}^{3}\sum_{j=0}^{3}\sum_{k=0}^{3}a_{ijk}x^{i}y^{j}z^{k} \tag{5.2}$$

Figure 5.5 shows the grid after 3D cubic interpolation based on the same grid in Figure 5.3.

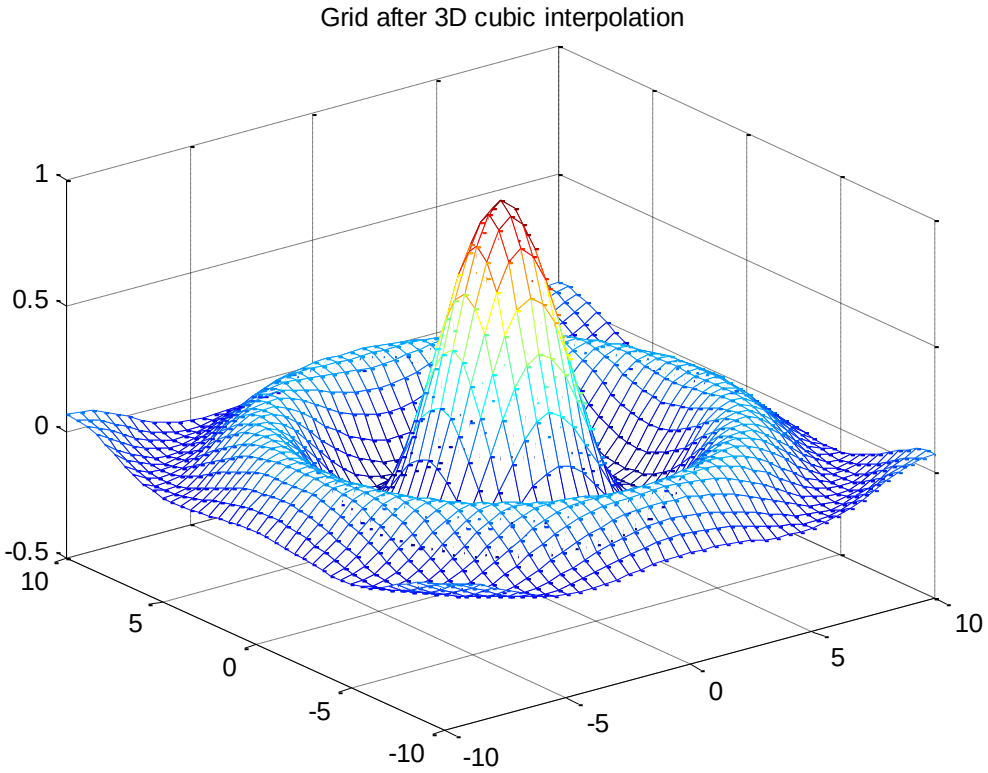


Figure 5.5 *3D cubic interpolation based on sample grid*

Similarly, we have 3D nearest neighbor interpolation which has similar principle in 2D. Figure 5.6 shows the example of 3D nearest neighbor interpolation based on grid of Figure 5.3

Among 3D linear, nearest and cubic interpolation methods, we can observe that cubic interpolation has the highest smoothness while the nearest neighbor has the worst smoothness.

There are also many other popular 3D interpolation methods. Natural neighbor interpolation was discussed in [3] in details. The extension from 2D to 3D was based on the vertices of the triangles. 3D barycentric interpolation, which is also called K nearest neighbor method and spherical spline interpolation were introduced in [22]. [23] proposed a fuzzy 3D interpolation model.

However, most of the literatures mentioned above are image processing related work, in WLAN-based positioning, the application of 3D interpolation and extrapolation is not widely used. Therefore, we should pay more attention to do research in this field.

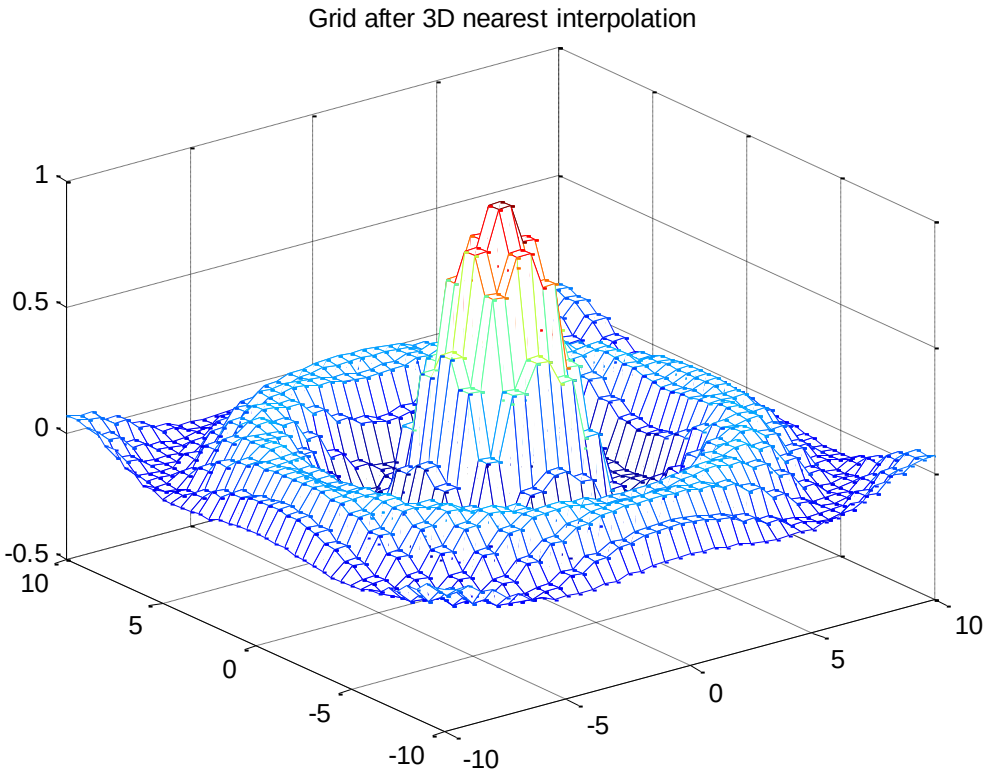


Figure 5.6 *3D nearest neighbor interpolation based on sample grid*

6. MEASUREMENT DATA ANALYSIS

The measurement data for simulation was collected with Windows tablet with HERE mapping tools, in 3 different building in Tampere University of Technology, Finland.

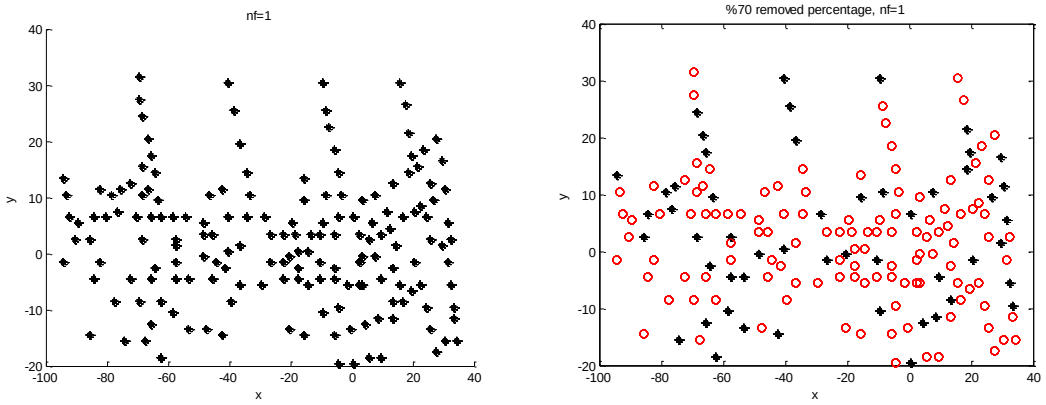
The information about the buildings are listed in Table 6.1

Table 6.1 *Measurement data information for three buildings*

Building	Number of Floor	Number of APs	Floor Height (constant for all floors)	Number of measurement points
A university building (Building 1)	4	309	3.7	1479
Same university building with renewed infrastructure (Building 2)	4	238	3.7	505
Another university building (Building 3)	3	354	3.5	584

To test the accuracy of the interpolation and extrapolation methods, a random percentage of grid points for each floor are extracted. The removed points are saved as reference points. Those remained points remained for interpolation. The RSS error is measured by the difference of RSS values between the reference point and interpolated point.

Figure 6.1 (a) shows the measurement points in first floor. Figure 6.1(b) shows the measurements with 70% points extracted. Star symbols refer to the points selected to do interpolation. Red circles represent the points remained as reference point.



(a) (b)

Figure 6.1 *Measurements in first floor of a university building and 70% removed percentage measurements at the same floor*

Delaunay Triangulation, natural neighbor, nearest neighbor, linear and cubic interpolation methods are applied in this simulation. For extrapolation, three methods are considered. The first is no extrapolation at all. The second is filling the NaN values with a sufficiently low value, namely -95 dBm, and the third one is to fill the NaN values with the minimum RSS value each AP heard at that floor.

Figures 6.2-6.4 shows the power maps of one building in campus generated by different interpolation and extrapolation methods. Figure 6.2 shows natural neighbor, linear, nearest neighbor and cubic interpolation in the condition without extrapolation respectively. Figure 6.3was generated with the extrapolation which filled the NaN values with -95 dBm. Figure 6.4 shows the plot with extrapolation that filled the NaN values with the minimum RSS value each AP heard at that floor.

In each figure, red circles represent the reference the reference point. Star symbols represent the selected points. All these figures are generated based on 10 percent removed from each floor. Because the measurements were extracted randomly each time, the positions of the interpolated points and reference points are different from Figures 6.2-6.4. However, in each plots, the number of interpolated points are 67, while the number of reference points are 7. Thus, the results are still reasonable to be used for comparison.

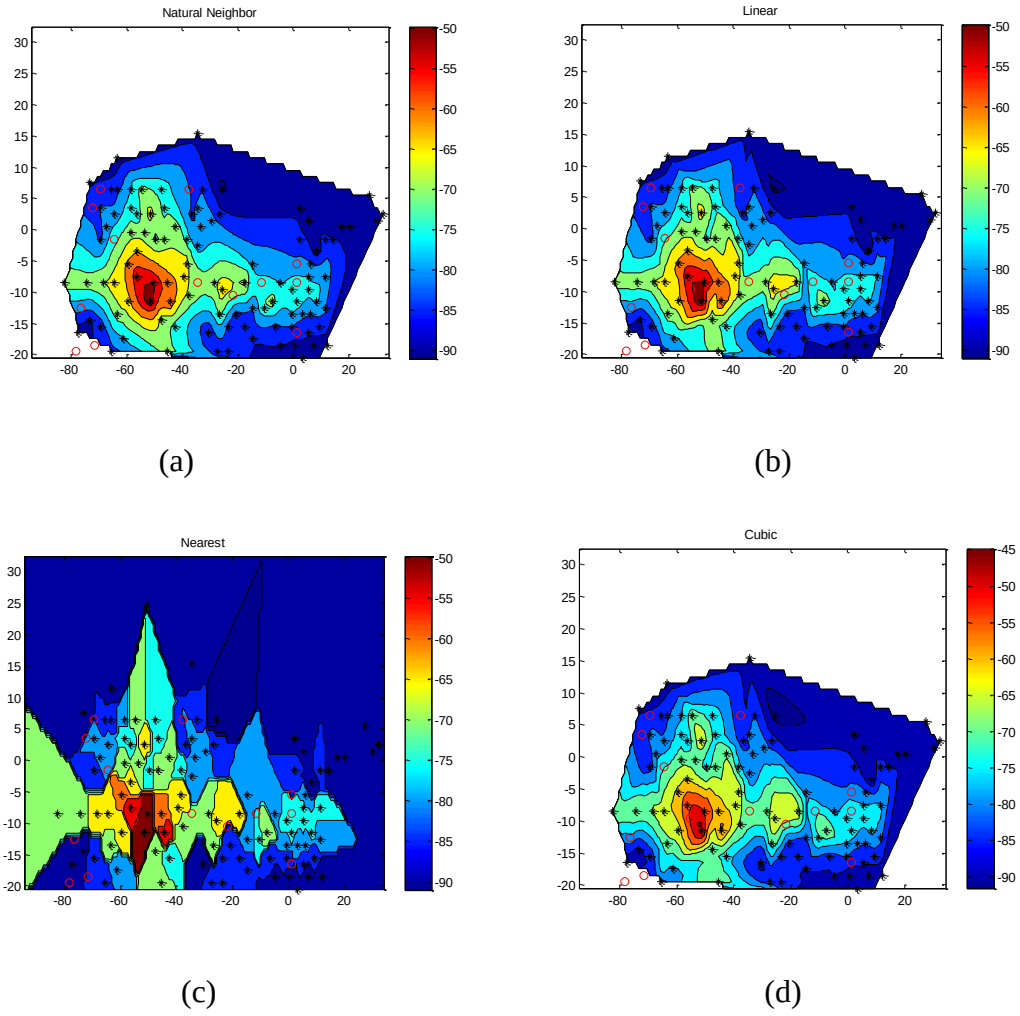
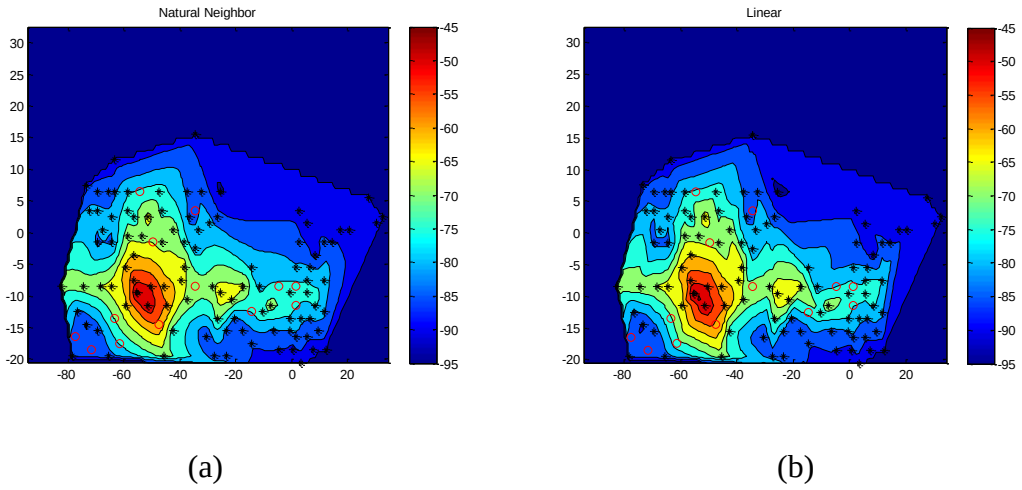
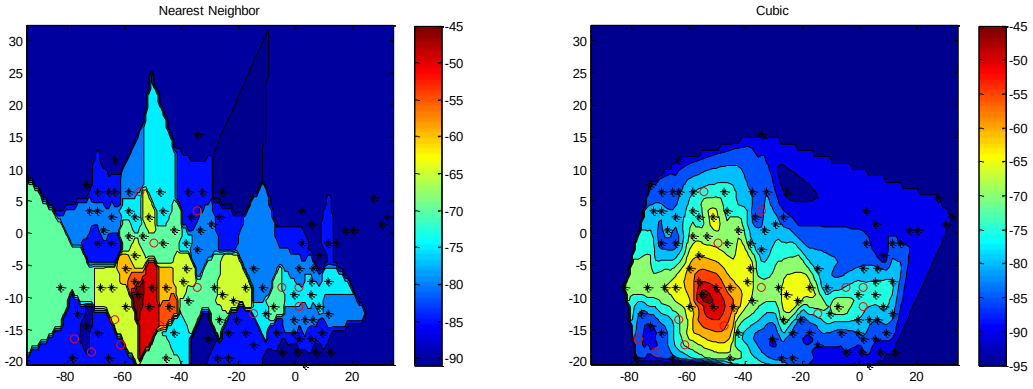


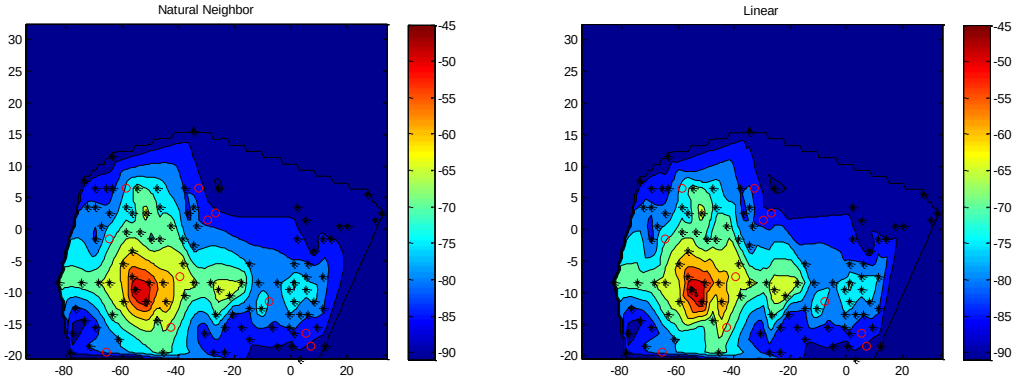
Figure 6.2 *Power maps without extrapolation: (a)=natural neighbor, (b)=linear, (c)=nearest neighbor, (d)=cubic*



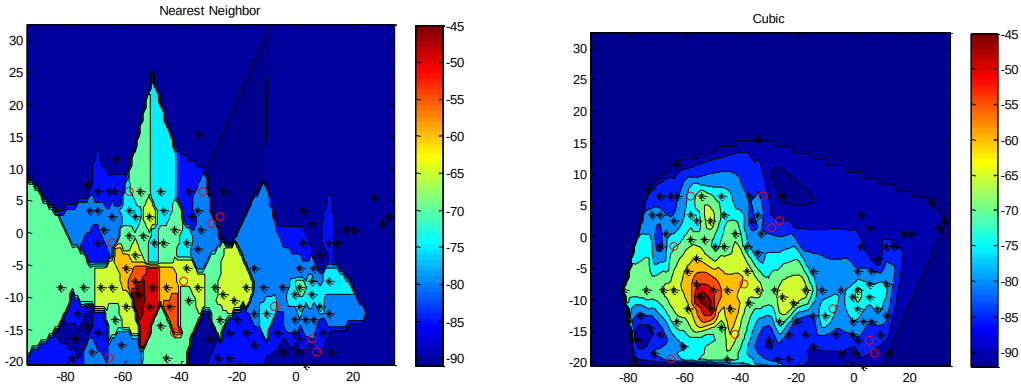


(c) (d)

Figure 6.3 *Power maps with extrapolation (Filling with -95): (a)=natural neighbor, (b)=linear, (c)=nearest neighbor, (d)=cubic*



(a) (b)



(c) (d)

Figure 6.4 *Power maps with extrapolation (Filling with minimum RSS value): (a)=natural neighbor, (b)=linear, (c)=nearest neighbor, (d)=cubic*

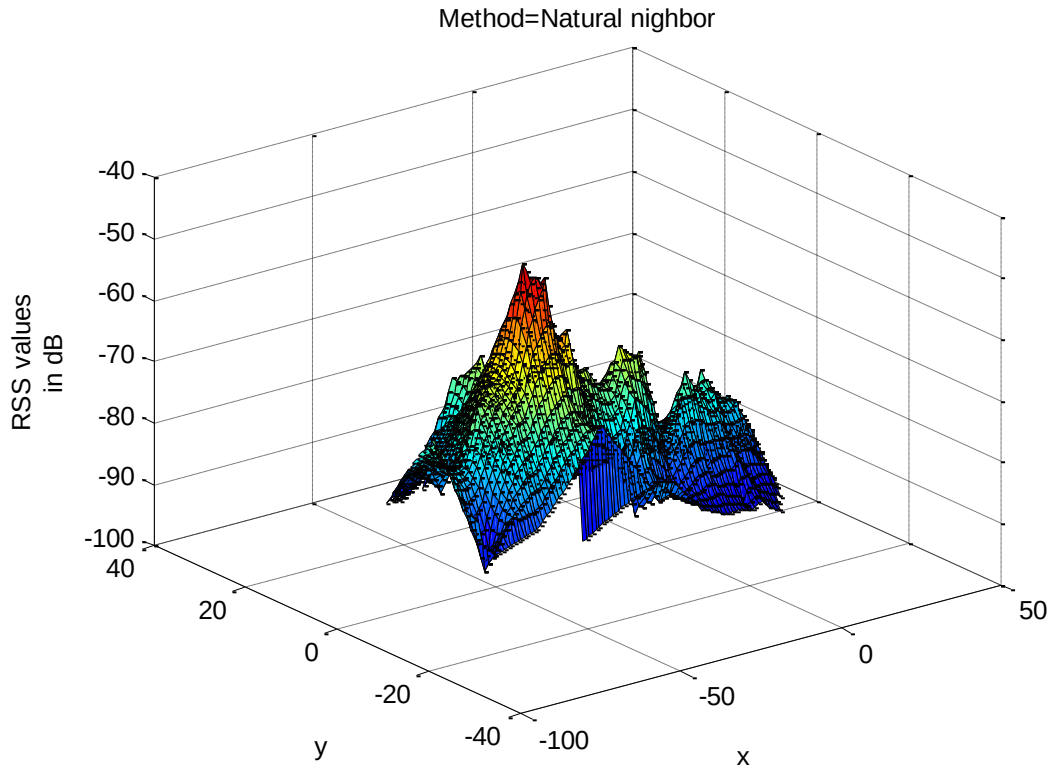
In Figure 6.2, we can observe that Figure 6.2 (a), Figure 6.2 (b) and Figure 6.2 (d) all have white area in the plots. This means that natural neighbor, linear and the cubic interpolation all have limits where no points are heard by that AP at white area. However, we can observe that nearest neighbor interpolation has different performance from other 3 methods. In Figure 6.2(c), no white area existed and this means that those points which are not heard by the AP are filled with RSS values, and these values are determined by the value of nearest point which heard by that AP. This is because there is no limit in nearest neighbor calculation. Nevertheless, there are limits in natural neighbor, linear and cubic interpolation. For example, for the linear interpolation, only points inside the triangles are used to determine the interpolant. In the nearest neighbor, no similar limits are required. This would have effect on the final results presented in the chapter 7.

In Figure 6.3, because of the extrapolation, the points which are not heard by that AP was filled with -95 dBm RSS values. Thus, white areas are filled with the color which corresponds to -95 dBm. In Figure 6.3 (a), Figure 6.3 (b), Figure 6.3 (c), Figure 6.3 (d), the white areas were filled with colors corresponding to -95 dBm and no white area existed any more.

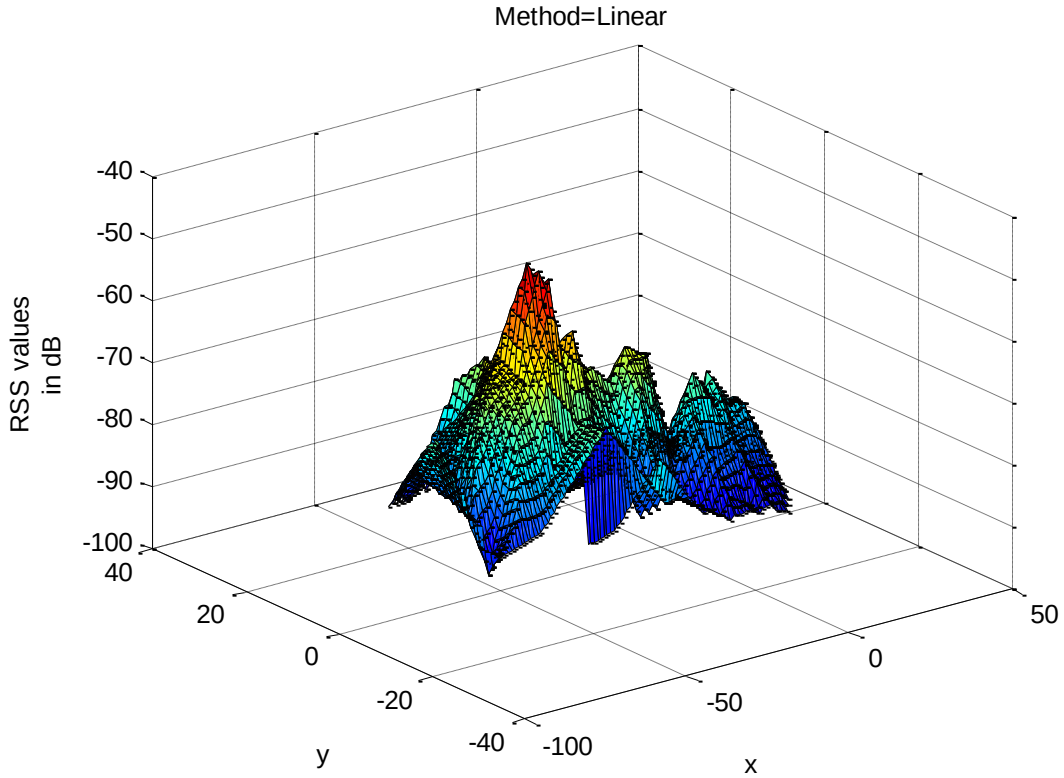
Similarly, in Figure 6.4, those points are not heard by AP was filled with the minimum values heard by that AP. Thus, in Figure 6.4 (a), Figure 6.4 (b), Figure 6.4 (c), Figure 6.4 (d), the white areas are filled with the value corresponding to the minimum RSS heard by that AP.

Moreover, we can notice the plots generated by cubic, linear and natural neighbor methods are smoother than the plot of nearest neighbor. The shape of Figure 6.4(d) is smoother than the shape in Figure 6.4 (c). This difference is more obvious in Figure 6.5. This figure was generated without extrapolation for the same floor same building.

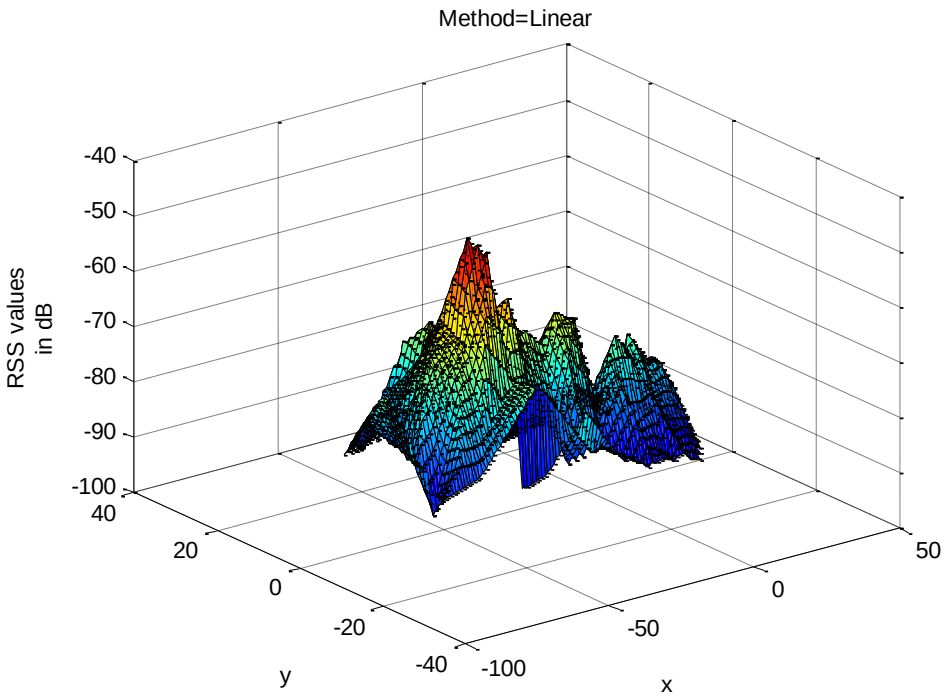
Figure 6.5 was generated based on 10% removed points from second floor in building1. From the figure, we can observe that Figure 6.5 (d) has best smoothness among other plots, while Figure 6.5 (c) has the worst smoothness. Thus, cubic interpolation has best smoothness while nearest neighbor has the worst shape.


Method=Natural nighbor
RSS values
in dB
-40
-50
-60
-70
-80
-90
-100
40
20
0
-20
-40
y
-100
-50
0
50
x

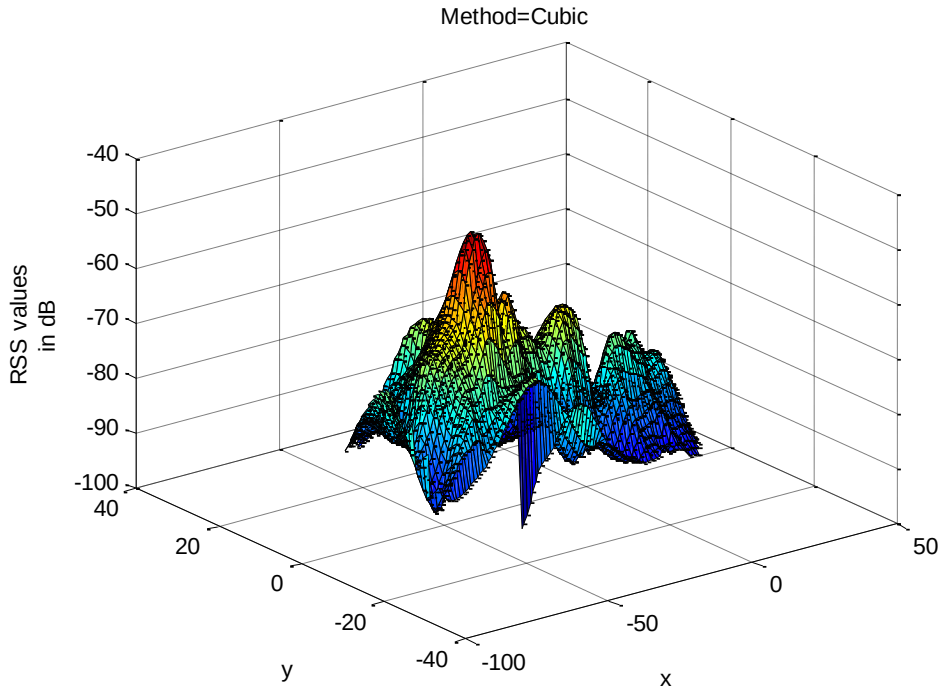
(a)


Method=Linear
RSS values
in dB
-40
-50
-60
-70
-80
-90
-100
40
20
0
-20
-40
y
-100
-50
0
50
x

(b)



(c)



(d)

Figure 6.5 *Surface plot of interpolation methods: (a)=natural neighbor, (b)=linear, (c)=nearest neighbor, (d)=cubic*

7. RESULTS AND DISCUSSION

In chapter7, firstly, the procedure of processing the data is explained. Then we present the measurement based results. Finally, the comparison and analysis between interpolation and extrapolation methods are done.

7.1 Measurement analysis procedure

The analysis procedure of this thesis is explained in Figure 7.1.

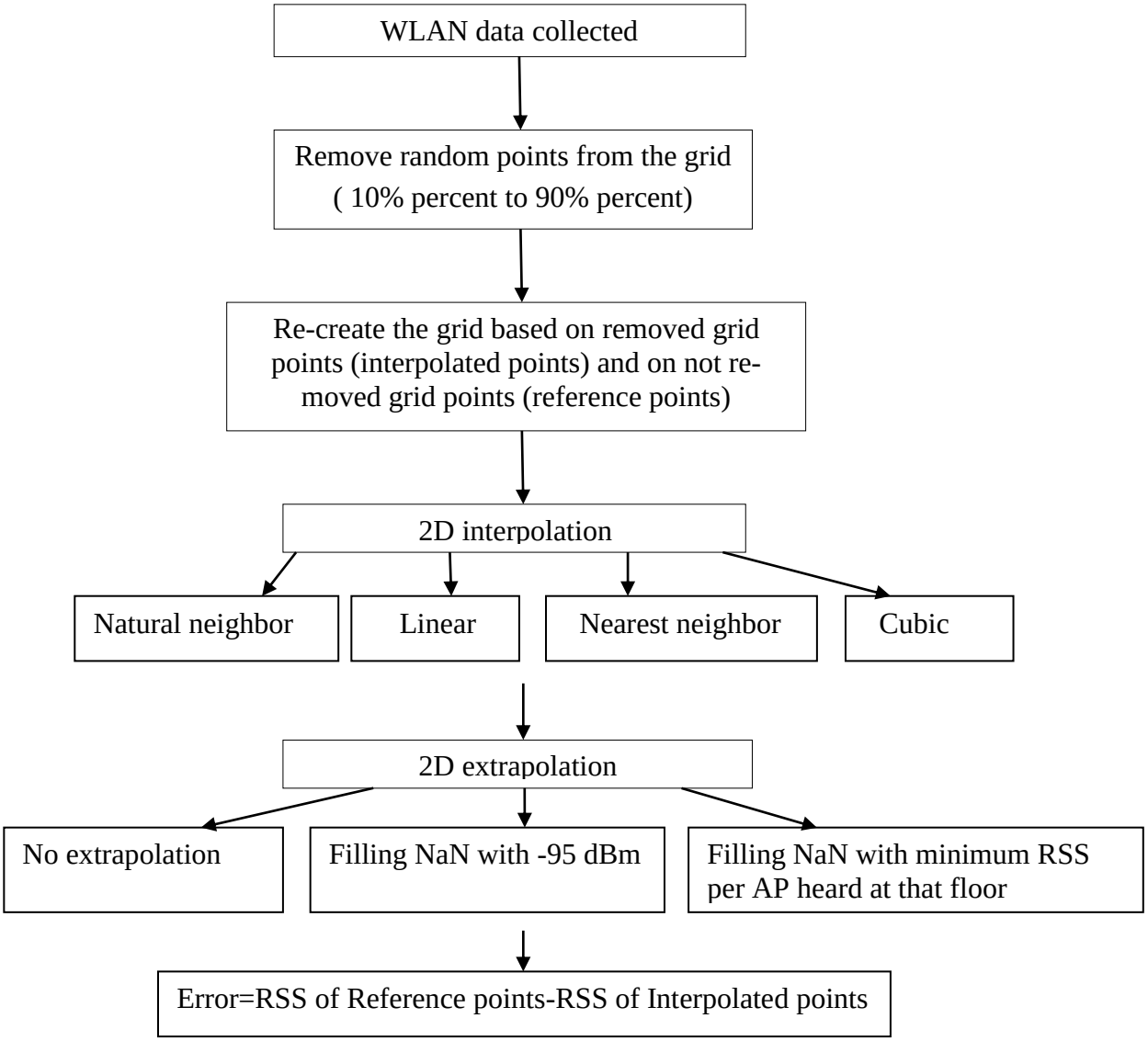
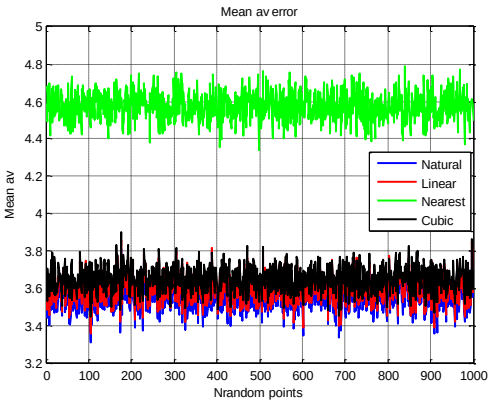


Figure 7.1 *Block Diagram of analysis procedure*

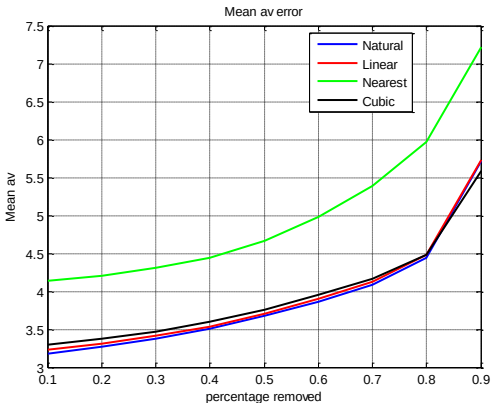
This procedure has done in 1000 times in order to obtain more accurate results.

7.2 Measurement based results

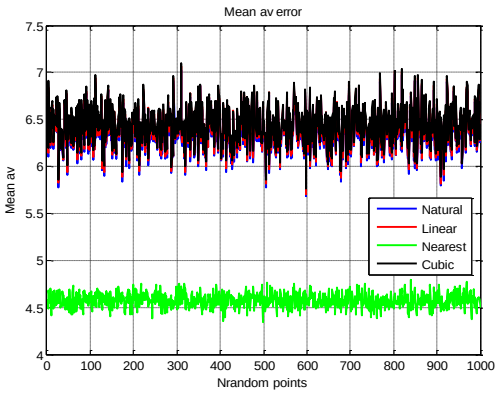
Figure 7.2 to Figure 7.4 shows the results for building1, building2 and building3 respectively. From figures 7.2-7.4, (a), (b) represent the results without extrapolation. (c), (d) show the condition filling NaN values with -95. (e), (f) show the results which filled with minimum value in each AP at that floor. The plots (a), (c), (e) show the average absolute error of 1000 operation times. Plots (b), (d), (f) indicates the mean absolute error from different removed percentage. These are generated based on taking the average value of 1000 times' results.



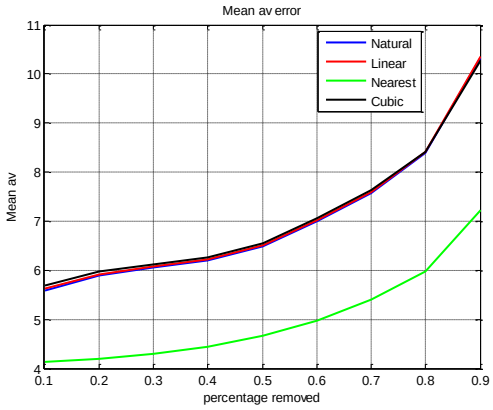
(a)



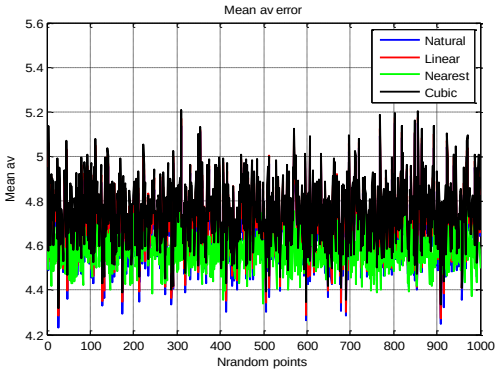
(b)



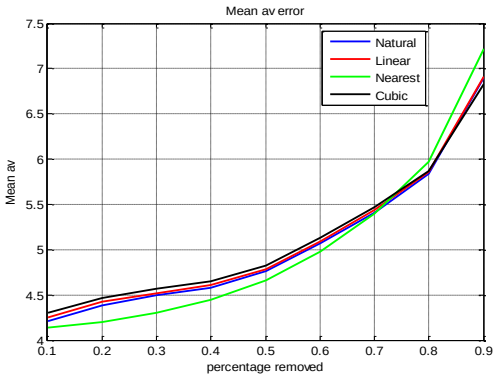
(c)



(d)

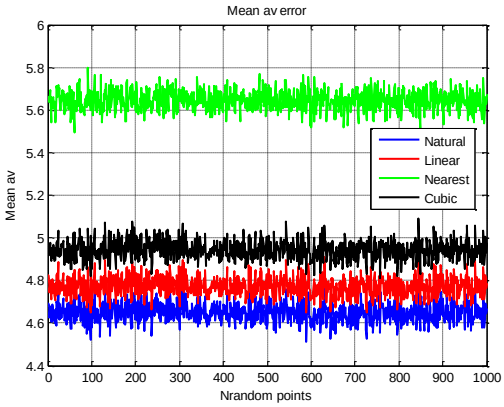


(e)

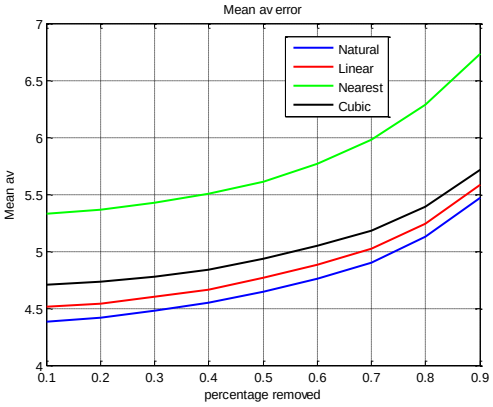


(f)

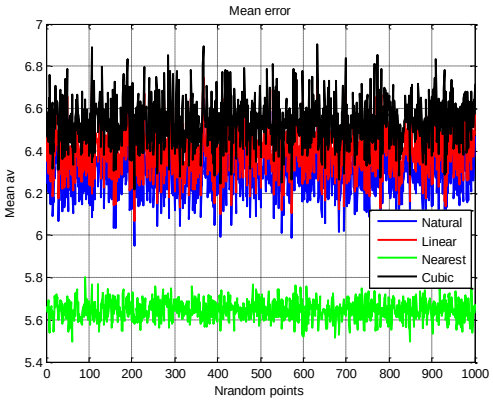
Figure 7.2 *Results for Building 1: (a)=Mean error of 1000 points with extrapolation, (b)=mean error for 4 interpolation methods without extrapolation, (c)= Mean error of 1000 points(filled NaN with -95dBm), (d)= mean error for 4 interpolation methods(filled NaN with -95dBm),(e)= Mean error of 1000 points(filled NaN with minimum RSS per AP), (f)= mean error for 4 interpolation methods(filled NaN with minimum RSS per AP)*



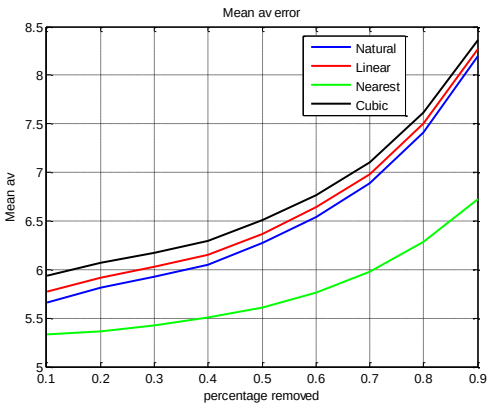
(a)



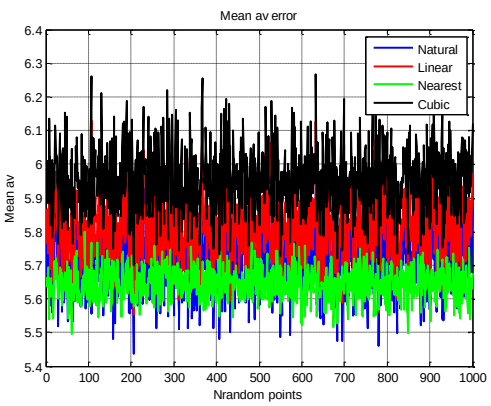
(b)



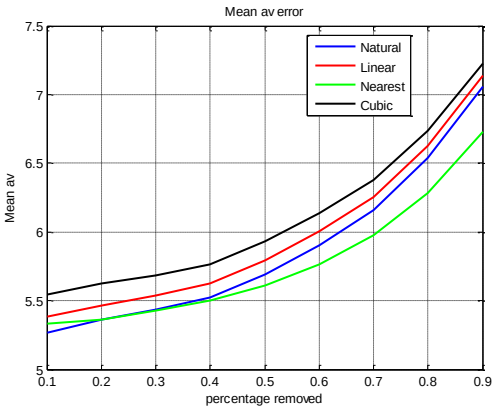
(c)



(d)

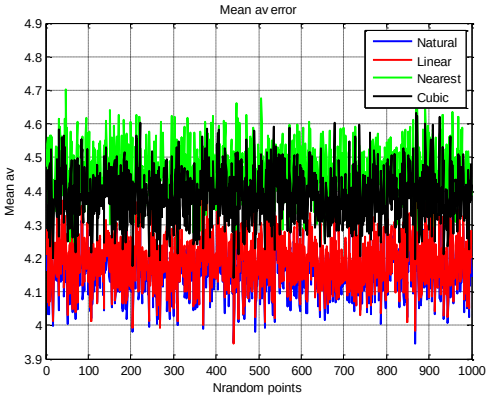


(e)

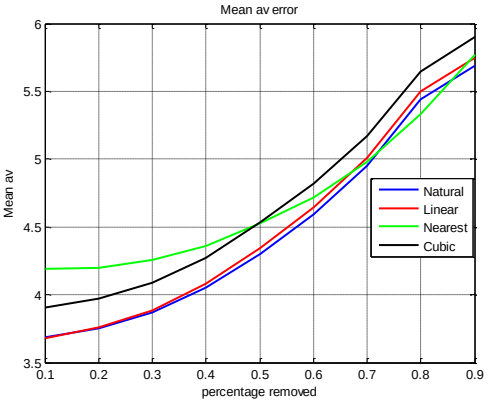


(f)

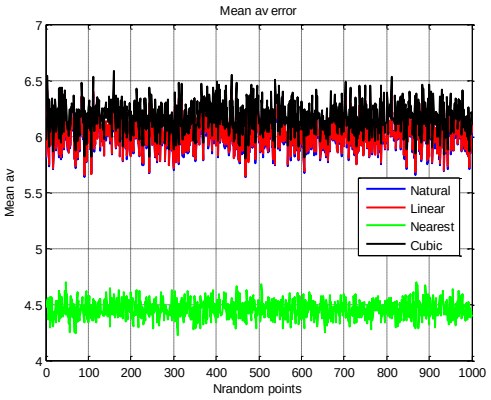
Figure 7.3 *Results for building2: (a)=Mean error of 1000 points with extrapolation, (b)=mean error for 4 interpolation methods without extrapolation, (c)= Mean error of 1000 points(filled NaN with -95dBm), (d)= mean error for 4 interpolation methods(filled NaN with -95dBm),(e)= Mean error of 1000 points(filled NaN with minimum RSS per AP), (f)= mean error for 4 interpolation methods(filled NaN with minimum RSS per AP)*



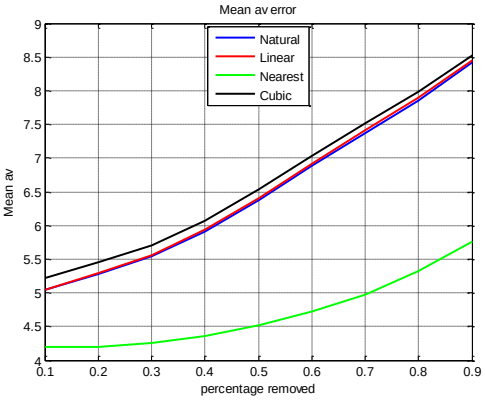
(a)



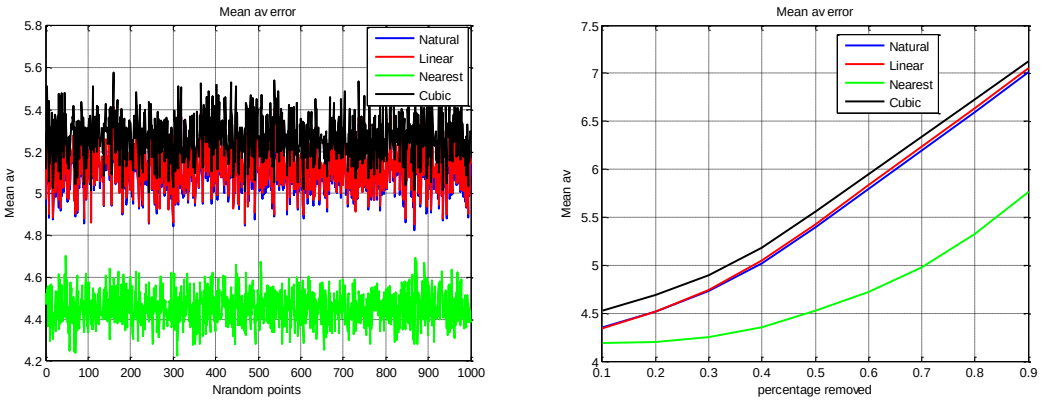
(b)



(c)



(d)



(e) (f)

Figure 7.4 *Results for building 3: (a)=Mean error of 1000 points with extrapolation, (b)=mean error for 4 interpolation methods without extrapolation, (c)= Mean error of 1000 points(filled NaN with -95dBm), (d)= mean error for 4 interpolation methods(filled NaN with -95dBm),(e)= Mean error of 1000 points(filled NaN with minimum RSS per AP), (f)= mean error for 4 interpolation methods(filled NaN with minimum RSS per AP)*

7.3 Analysis and discussion

Basically, during the step of removing some grid points, the points are extracted randomly. Therefore, for each operation time, the error was not the same. It makes sense that the error fluctuated, as seen in plot (a), (c), (e) of Figures 7.2-7.4.

Generally, without extrapolation, as the comparison between four interpolation methods showed in Figure 7.2-7.4 (b), the nearest neighbor interpolation has obvious higher error than the error of natural neighbor, linear and the cubic interpolation. The performance of natural neighbor, linear and the cubic interpolation are similar. However, differences can be seen still such as natural neighbor has better performance than linear and cubic interpolation. Without extrapolation, the error from low to high is given by: natural neighbor, linear, cubic, nearest neighbor respectively.

With the extrapolation which fills the missing values with -95 dBm, as the comparison between four interpolation methods showed in Figure 7.2-7.4 (d), the green line which represents nearest neighbor interpolation did not change compared to the green line in Figure 7.2-7.4 (b). This result is consistent with the power maps presented in chapter 6, Figure 6.2 (c) and Figure 6.3(c). Moreover, in Figures 7.2(d), 7.3(d) and 7.4 (d), we can observe that: the error of blue, red, black line which represent natural neighbor, linear and the cubic interpolation respectively were increased compared to the error showed in

Figure 7.2-7.4 (b). In Figure 7.2-7.4 (d), the error from low to high is given by: the nearest neighbor, natural neighbor, linear, and the cubic interpolation respectively.

With the extrapolation which fills the missing values with the minimum values from each AP, as the comparison between four interpolation methods showed in Figure 7.2-7.4 (f), errors are decreased due to more accurate extrapolation compared to the error showed in Figure 7.2-7.4 (d) . However, in Figure 7.2-7.4 (f), the error from low to high is still given by: the nearest neighbor, natural neighbor, linear, and cubic interpolation, respectively.

To put it more straightforward, Table 7.1 shows the mean error of each interpolation and extrapolation method for the three considered buildings. The lowest error in each case is shown in black boldfaced in numbers.

The comparison of interpolation methods of other respect such as execution time, complexity, memory occupation and smoothness are discussed in Table 7.2.

Thus, according to Table 7.1, without extrapolation, we can see that, natural neighbor interpolation has best performance. The error from low to high is given by: the natural neighbor, linear, cubic, and the nearest neighbor respectively. However, we can observe that the nearest neighbor interpolation has best performance while extrapolation was implemented in the measurement. With extrapolation, the error from low to high is given by: nearest neighbor, natural neighbor, linear, and the cubic interpolation respectively. Therefore, based on the analysis above, the nearest neighbor interpolation with extrapolation which filled missing values with minimum RSS heard by AP would have lowest error in the measurement.

One point that has to be mentioned is that with accurate extrapolation methods, the average error of the natural neighbor, linear and cubic interpolation can be decreased. For example, in the results of building1, the error of natural neighbor with extrapolation which filled missing values with -95dBm was 6.3725 dB, and the error with extrapolation which filled missing values with minimum RSS heard by that AP was 4.6803 dB. Thus, we can expect that with more accurate extrapolation methods applied, the measurement errors of the natural neighbor, linear and the cubic interpolation can be decreased and even have better performance than the nearest neighbor interpolation.

In practical, the measurements we collected from the building are restricted because there might be some area cannot be accessed. However, to obtain more accurate simulation, a large number of available measurements are required. Thus, extrapolation is useful in WLAN-based positioning. When low accuracy extrapolation methods are applied, nearest neighbor interpolation is suggested due to its easy implementation and low measurement error. When more accurate extrapolation can be used, the natural neighbor interpolation can be considered to obtain more accurate measurement results.

Table 7.1 *Mean error for interpolation and extrapolation for average percent removed points*

		Error for Natural neighbor (dB)	Error for linear interpolation (dB)	Error for Nearest neighbor(dB)	Error for cubic interpolation(dB)
Building1	No extrapolation	**3.5592**	3.5993	4.5748	3.6519
	Fill with -95 dBm	6.3725	6.4035	**4.5748**	6.4451
	Fill with minimum value	4.6803	4.7114	**4.5784**	4.7529
Building2	No extrapolation	**4.6499**	4.7715	5.6445	4.9420
	Fill with -95 dBm	6.2947	6.3959	**5.6445**	6.5350
	Fill with minimum value	5.7137	5.8149	**5.6445**	5.9540
Building3	No extrapolation	**4.1661**	4.1904	4.4556	4.3862
	Fill with -95 dBm	6.0106	6.0287	**4.4556**	6.1722
	Fill with minimum value	5.1084	5.1265	**4.4556**	5.2700

Table 7.2 *Comparison between interpolation methods*

	Complexity	Execution time	Memory occupation	Smoothness
Natural neighbor	High	Long	Middle	Good
Linear	Low	Long	Middle	Good
Nearest neighbor	Lowest	Quick	Low	Bad
Cubic	Highest	Longest	High	Excellent

8. CONCLUSION AND FUTURE WORK

8.1 Conclusion

The main contributions of this thesis have been to analyze and to make a comparison between different interpolation and extrapolation methods based on WLAN-based positioning. The author studied several popular WLAN-based positioning methods. RSS-based fingerprinting and path-loss models were overviewed. Then, various interpolation and extrapolation methods were presented. Moreover, the applications of those methods in WLAN context were discussed. This thesis focused on triangulation-based methods, namely the natural neighbor, the linear, the nearest neighbor and the cubic interpolation. Due to the complexity and uncertainty, only two simple extrapolation methods are studied, namely padding with a constant small value and padding with variable (minimum per AP) values. Further, the extension of Delaunay triangulation in 3 dimension space was discussed. The principles of cubic and linear interpolation methods in 3D were also introduced.

Regarding the measurement-based results, the nearest neighbor interpolation has the ability to fill unknown points with nearest values. Thus, the extrapolation has little effect on it. The implementation of this method is simple, and the execution time is quick. However, the smoothness of this method is the worst among the methods tested in this thesis. The reason for nearest neighbor interpolation to have less error than the other methods studied in this thesis is because the extrapolation is not exact enough. With more complex and accurate extrapolation, the other three methods would perform better. For the other three methods, the performance from low level to high is cubic, linear, natural neighbor.

8.2 Future work

There is still room for improving this thesis. The work will be continued by implementing more complex interpolation and extrapolation methods, such as spline, ordinary kriging, inverse-distance weighted, adaptive normalized convolution interpolation and linear extrapolation methods. Also, more details of the comparison should be made. For example, the degree of complexity, the exact execution time, the numerical data for the memory occupation and so forth. Also, the impact on the positioning accuracy is still an open issue.

Moreover, during the measurement procedure, a limitation was set that only more than

10 points heard by each AP has the obligation to participate in calculation. In the future, the number can be reduced to 5 or 6 to study more realistic occasions.

Finally, the requirements to implement interpolation and extrapolation in 3-dimensional space are increased in the research field. The study of those models applied in 3D should be studied in the future.

REFERENCES

[1] S. Shrestha, "RSS-based position estimation in cellular and WLAN networks", Master of Science thesis, Tampere University of Technology. January, 2012.

[2] A.G. Dempster, "Dilution of precision in angle-of-arrival positioning systems", in *Electronics letters*, vol. 42, no. 5, March 2006

[3] L.Liang and D.Hale, "A stable and fast implementation of natural neighbor interpolation". April, 2010. [online].Available: http://www.cwp.mines.edu/Meetings/Project10/cwp-657.pdf. [Accessed: Feb 14, 2015]

[4] P. K. Zadeh, "Investigations of Dempster-Shafer theory in the context of WLAN-based indoor localization", Master of Science thesis, Tampere University of Technology. June, 2013.

[5] E.Laitinen *et al.*, "Access Point Significance Measures in WLAN-based location", in *Positioning Navigation and Communication (WPNC)*, March 2012, pp. 24-29

[6] Yi.Li *et al.*, "The Application of Natural Neighbor Interpolation in Real-time Environments", in *Cyberworlds, 2008 international conference*, pp. 306-313, Sept. 2008

[7] S.Üreten *et al.*, "Interference Map Generation Based on Delaunay Triangulation in Cognitive Radio Networks", IEEE 13th *International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 134-138, 2012,

[8] Zhongxue Li *et al.*, "Improvement on Inverse Distance Weighted Interpolation for Ore Reserve Estimation", *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference*, pp. 1703-1706, 2010

[9] J.Bozek *et al.*, "Comparative Analysis of Interpolation Methods for Bilateral Asymmetry", in *Elmar*, pp.1-7, Sept.2010

[10] ChunlinWu *et al.*, "Fast data extrapolating", *Journal of Computational and Applied Mathematics*, pp. 146-157, 2007

[11] K.Jbilou and H.Sadok, "Analysis of some vector extrapolation methods for solving systems of linear equations", in *Numer.Math*, page 74, 1995

[12] K. Jbilou and H. Sadok, “Vector extrapolation methods. Applications and numerical comparison”, in *Journal of Computational and Applied Mathematics*, pp.149-165, 2000

[13] M.Natale *et al.*, “A Novel Gaussian Extrapolation Approach for 2D Gel Electrophoresis Saturated Protein Spots”, in *Genomics Proteomics Bioinformatics*, pp.336 – 344, 2012

[14] E. E. Mantilla *et al.*, “IEEE 802.11 b and g WLAN Propagation Model using Power Density Measurements at ESPOL”, in *World Academy of Science, Engineering and Technology*, 2010

[15] B. Li .Author, “Using two global positioning system satellites to improve wireless fidelity positioning accuracy in urban canyons”, in *IET Commun.*, vol. 5, pp. 163–171, Jan. 2011

[16] W.V. Belle, Antenna Finding and Interpolation/Extrapolation of Signal Strength. November ,2006 .[online].Available: http://werner.yellowcouch.org/Papers/antenna/ [Accessed: Feb 14, 2015]

[17] K.Arai and H.Tolle, “Color Radiomap Interpolation for Efficient Fingerprint WiFi-based Indoor Location Estimation”, in *International Journal of Advanced Research in Artificial Intelligence*, vol. 2, no.3, 2013

[18] C. Andersson, Solving linear equations on parallel distributed memory architectures by extrapolation. 2000.[online].Available: http://www.nada.kth.se/~christe/masters-thesis.pdf [Accessed: Feb 14, 2015]

[19] Ran Tao *et al.*, “A Comparison of US versus MR Based 3D Prostate Shapes Using Radial Basis Function Interpolation and Statistical Shape Models”, in *Journal of Biomedical and Health Informatics*, May, 2014

[20] Visualization, Summer Term 03, VIS, University of Stuttgart, 3. Interpolation and Filtering. 2003. [online]. Available: https://graphics.ethz.ch/teaching/former/scivis_07/Notes/stuff/StuttgartCourse/VIS-Slides-03-Interpolation_and_Filtering.pdf [Accessed: Feb 14, 2015]

[21] V. Sacristán, Delaunay triangulation of point sets.2013.[online]. Available: http://www-ma2.upc.es/vera/wp-content/uploads/2012/10/Delaunay.pdf [Accessed: Feb 14, 2015]

[22] I. Nouira *et al.*, “EEG Potential mapping by 3D interpolation methods”, in *Multimedia Computing and Systems (ICMCS)*, pp.469-474, 2014

[23] N.Bizon and M.Stork, "2D Image pre-Processing - Part I: Fuzzy 3D Interpolation", in *Applied Electronics (AE)*, 2011 International Conference on, pp. 1-4, Sept. 2011

[24] W. K. Jenkins *et al.*, "Nearest neighbor and generalized inverse distance interpolation for fourier domain image reconstruction", in *Acoustics, Speech, and Signal Processing*,1985

[25] S.Woo *et al.*, "Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR", in *Automation in Construction*, pp. 3-13, 2010

[26] M.Hossain *et al.*, "SSD: A Robust RF Location Fingerprint Addressing Mobile Devices' Heterogeneity", in *Transactions on mobile computing*, vol. 12, no. 1, Jan.2013

[27] Landu Jiang, "A WLAN fingerprinting based indoor localization technique", Master of science thesis, The Graduate College at the University of Nebraska, July, 2012

[28] S. Shrestha *et al.*, "Deconvolution-based indoor localization with WLAN signals and unknown access point locations", in *Localization and GNSS (ICL-GNSS)*, pp.1-6, June 2013

[29] J. Talvitie *et al.*, "Access Point Significance Measures in WLAN-based Location", in *Localization and GNSS (ICL-GNSS)*, pp.1-6, June 2014

[30] S.Shrestha *et al.*, "RSSI channel effects in cellular and WLAN Positioning", in *Positioning Navigation and Communication (WPNC)*, pp.187-192, March 2012

[31] S. Mazuelas *et al.*, "Robust Indoor Positioning Provided by Real-Time RSSI Values in Unmodified WLAN Networks", in *journal of selected topics in signal processing*, vol. 3, no. 5, October 2009

[32] Xuan Yang and Jihong Pei, "3D Interpolation of Image Elastic Deformation Using Delaunay Triangulation", in *Bioinformatics and Biomedical Engineering*, pp.1-4, June 2009

[33] Xiao Shu, Bicubic interpolation. March 25, 2013. [online].Available: http://www.ece.mcmaster.ca/~xwu/3sk3/interpolation.pdf [Accessed: Feb 14, 2015]

[34] M.P. Foster and A.N. Evans, "An Evaluation of Interpolation Techniques for Reconstructing Ionospheric TEC Maps", in *Geoscience and Remote Sensing*, pp.2163-2154, July 2008